

83779

STIC-Biot ch/ChemLib

Fr m: Chan, Christina
S nt: Tuesday, January 07, 2003 1:14 PM
T : Duffy, Patricia; STIC-Biotech/ChemLib
Subject: RE: Please Rush -- Amendment Due this Biweek

Imp rtance: High

Pl ase rush. Thanks Chris

Chris Chan
TC 1600 New Hire Training Coordinator and SPE 1644
308-3973
CM-1, 9B19

-----Original Message-----

From: **Duffy, Patricia**
Sent: Tuesday, January 07, 2003 1:12 PM
To: Chan, Christina
Subject: Please Rush -- Amendment Due this Biweek
Importance: High

In re: 09/900,766

Please search SEQ ID NOs:1 and 2.
Please perform an interference search.
Please print out top 50 hits in each category.
Thanks mucho.

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Patricia A. Duffy
CM1-8D05
AU 1645
703-305-7555

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 1/8/03
Date Completed: 1/18/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: ✓
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: PK
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:52:01 ; Search time 49.0077 Seconds
(without alignments)
1827.149 Million cell updates/sec

Title: US-09-900-766-1
Perfect score: 3522
Sequence: 1 EVQLQQSGDPLVKPGASVKI.....EATHKSTSTPIVKSFNRNES 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1644	46.7	573	22 AAG64459	Mouse 6D9 catalyti
2	1113	31.6	245	18 AAW33375	Staphylococcus ent
3	1107	31.4	257	22 AAU14103	Peptide sequence f
4	1096.5	31.1	720	22 AAG65572	Amino acid sequenc
5	1093	31.0	230	22 AAB67339	Staphylococcus aur
6	1091	31.0	245	18 AAW35374	Staphylococcus ent
7	1069	30.4	592	22 AAB83838	Amino acid sequenc
8	1069	30.4	595	22 AAW86003	Anti-5T4 single ch
9	1059.5	30.1	464	16 AAR76088	MAB 55.1 heavy cha
10	1055.5	30.0	445	16 AAR76085	MAB 55.1 heavy cha

11	1048	29.8	230	14 AAR45012	Staphylococcal ent
12	1044	29.6	230	23 ABB76235	Staphylococcus aur
13	1035	29.4	230	12 AAB13204	Staphylococcal ent
14	1028	29.2	597	23 AAE18377	Human N-terminal D
15	1025.5	29.1	464	19 AAW83041	Anti-Fas MAB HFE7A
16	1025.5	29.1	464	21 AAB14747	Mouse anti-Fas ant
17	1025.5	29.1	464	21 AAW90897	Murine anti-Fas an
18	1025.5	29.1	464	23 ABB74866	Humanised anti-Fas
19	1025.5	29.1	464	23 ABB74912	Humanised anti-Fas
20	1024.5	29.1	613	23 AAE18380	Human N-terminal D
21	1023.5	29.1	456	23 AAE18370	Human penton base
22	1023.5	29.1	493	23 AAE18379	Human N-terminal D
23	1023.5	29.1	510	23 AAE18378	Human N-terminal D
24	1022.5	29.0	438	23 AAE18372	Human penton base
25	1018	28.9	206	20 AAY39452	Antibody ABX-CBL 1
26	1018	28.9	465	16 AAR66758	Anti-tobacco mosai
27	980.5	27.8	711	20 AAW85692	MOTAbII fusion pro
28	977.5	27.8	626	21 AAY55081	Single chain Fv pr
29	972.5	27.6	243	20 AAY42294	Anti-5T4 secreted
30	972.5	27.6	243	20 AAY27407	5T4 scFv antibody
31	972.5	27.6	243	20 AAW86002	Murine anti-5T4 an
32	972.5	27.6	243	22 AAB83835	Amino acid sequenc
33	970	27.5	214	20 AAY44176	MAB Fab1365 heavy
34	967.5	27.5	488	20 AAW86004	Human B7-1.5T4.1 p
35	967.5	27.5	488	22 AAB83836	Amino acid sequenc
36	966	27.4	468	12 AAR13061	Monoclonal antibody
37	961.5	27.3	225	14 AAR40385	Monoclonal antibody
38	960	27.3	212	17 AAW15932	Antibody 7G12 heav
39	960	27.3	233	12 AAR13203	Staphylococcal ent
40	953.5	27.1	223	16 AAR75456	Mouse antibody FB3
41	953	27.1	220	15 AAR53802	FAB light chain fo
42	951.5	27.0	469	14 AAR40384	Monoclonal antibody
43	948	26.9	214	21 AAU78253	Mouse agglutinatio
44	948	26.9	257	22 AAU14104	Peptide sequence f
45	944.5	26.8	239	16 AAR76087	MAB 55.1 light cha

ALIGNMENTS

RESULT 1
AAG64459
ID AAG64459 standard; Protein; 573 AA.
XX
AC AAG644459;
XX
DT 21-SEP-2001 (first entry)
XX
DE Mouse 6D9 catalytic antibody.
XX
KW Mouse; 6D9; catalytic antibody; esterase activity.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT Misc-difference 28 /note= "Encoded by GCCTGA"
FT FT Misc-difference 35 /label= unknown
FT FT Misc-difference 46 /note= "Encoded by TNT"
FT FT Misc-difference 329 /note= "Encoded by GAGTAA"
FT FT Misc-difference 572 /note= "Encoded by GAGTAA"
FT FT Misc-difference 572 /note= "Encoded by TGTAA"
XX JP2001128681-A.
XX 15-MAY-2001.
XX 04-NOV-1999; 99JP-0314259.
XX

XX	Key	Location/Qualifiers	
FT	Misc-difference 20	/note= "can be mutated at this position"	
FT	Misc-difference 21	/note= "can be mutated at this position"	
FT	Misc-difference 24	/note= "can be mutated at this position"	
FT	Misc-difference 27	/note= "can be mutated at this position"	
XX	WO9736932-A1.		
PN			
XX	09-OCT-1997.		
XX			
PF	26-MAR-1997; 97WO-SE00537.		
PR	12-AUG-1996; 96US-0695692.		
PR	29-MAR-1996; 96SE-0001245.		
XX	(PHAA) PHARMACIA & UPJOHN AB.		
XX	Abrahmsen L, Antonsson P, Bjoerk P, Dohlsten M;		
PI	Forsberg G, Hansson J, Kalland T;		
XX			
DR	WPI; 1997-503052/46.		
XX			
XX	Conjugate of target seeking moiety and modified superantigen -		
PT	useful for activating the immune system to treat cancer, viral		
PT	infections, parasitic infestations and autoimmune diseases		
XX			
PS	Claim 4; Pages 38-39; 58pp; English.		
XX			
CC	This is the wild-type Staphylococcus enterotoxin SEE superantigen. This		
CC	SEE superantigen can be modified to be used in a novel conjugate. The		
CC	novel conjugate comprises a target seeking moiety and a modified wild		
CC	type superantigen. The modified superantigen retains its ability to		
CC	activate a subset of T cells, even though 1 or more wild-type amino acid		
CC	residues in at least 1 region which functions in determining binding to		
CC	T cell receptor (TCR) and activation of a subset of T cells has/have been		
CC	replaced. Such a modified superantigen can optionally be used as part of		
CC	a conjugate with a target seeking moiety, for activating the immune		
CC	system to treat a mammalian disease. A pharmaceutical composition can be		
CC	prepared comprising a modified antibody (preferably a Fab fragment fused		
CC	to a peptide moiety providing activation of T cells in Vbeta specific		
CC	manner) in which cysteines providing for interchain cysteine linkages in		
CC	the native antibody have been replaced (preferably by serine residues) to		
CC	prohibit cysteine formation. The modified wild-type superantigen is used		
CC	for treating cancer, viral infections, parasitic infestations and		
CC	autoimmune disease. The modified wild type superantigen has a lower		
CC	immunogenicity and reactivity with neutralising antibodies and has fewer		
CC	side-effects when used as a drug, compared to wild type superantigen.		
XX			
SQ	Sequence 245 AA;		
Query Match 31.0%; Score 1091; DB 18; Length 245;			
Best Local Similarity 85.3%; Pred. No. 5.2e-58;			
Matches 209; Conservative 9; Mismatches 15; Indels 12; Gaps 1;			
QY	226	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSRAITSSEKSAOFLTNTLLFGFTFG 285	
DB	1	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSRAITSSEKSAOFLTNTLLFGFTFG 60	
QY	286	HPWYNLLVDLGLSTAASTSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 345	
DB	61	HPWYNLLVDLGLSKGATNKKYGGKVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120	
QY	346	EKKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDQARHYLHGKFGLYNSDSFGGKVQ 405	
DB	121	EKKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDQARHYLHGKFGLYNSDSFGGKVQ 180	
QY	406	RGLIVFHSSEGSTSYDLFDAQQGYPTDLLRIYRNKNTINSENHLHIDLYTINSENHLIDL 240	

```
Matches 254; Conservative 20; Mismatches 60; Indels 336; Gaps 9;
QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYIMHWKSPGKLEWIGRINPNNGVTLY 60
DB 20 EVLOQSGDPLVKPGASVKISCKASGYFTGYIMHWKSPGKLEWIGRINPNNGVTLY 79
QY 61 NQKFKDKATLTVDKSTTAYMELRLSLTSDSAVYYCARSTMITNTYMDYWGQGTSTVTVSS 120
DB 80 NQKFKDKATLTVDKSTTAYMELRLSLTSDSAVYYCARSTMITNTYMDYWGQGTSTVTVSS- 138
QY 121 AKTTPPSVYPLAPGSAQAQTNSMTGLCLVKGYFPEPVTVTVNSGSLSSGVHTFPAVLQSD 180
DB 139 -----SGG----- 138
QY 181 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDSGGPSEKSEINEKDLRKK 240
DB 139 -----SGG----- 141
QY 241 SELOGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTGHPWYNLLVDLGSTA 300
DB 142 ----- 141
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEEKVPIINLWIDGKQ 360
DB 142 ----- 151
QY 361 TTVPIDKVKTSKKEVTVOELDLQARHLYLHGKFLYNSDSFGCKVQVGLIVFHSSEGSTVS 420
DB 152 -----GG----- 153
QY 421 YDLFDAQGOYPTDLLRIYRDNTTISTSLISLYLYTTSIVMTQPTSLVLSAGDRVTIT 480
DB 154 -----SSIVMTQPTSLVLSAGDRVTIT 176
QY 481 CKASQSVNDVAYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTFTLTISVQAED 540
DB 177 CKASQSVNDVAYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTFTLTISVQAED 236
QY 541 AAVYFCQDYNPPPTFGGKLEIKRADA-APTVSTFFPPSPBOLTSGGASVVCFLNNFYP 599
DB 237 LAVYFCQDYNPPPTFGGKLEIKRASKTGPSPFLAPSSKSTSGGTAALGCLVKDYFP 296
QY 600 KDINKYK-----IDGSEKQNGVLNSWTDQDSKDSYMSSTLTTLTKDEYERHNSYTCEAT 655
DB 297 EPTVTSWNSGALTSVGHVTPPAVLQS-----SGLYSLSVVTVPSLSLGT-OTVICNVN 348
QY 656 HKTSTSPVK 665
DB 349 HKPSNTKVDK 358

RESULT 8
AAW86003
ID AAW86003 standard; Protein; 595 AA.
XX
AC AAW86003;
XX
DT 15-MAR-1999 (first entry)
XX
DE Anti-5T4 single chain antibody 5T4Sabl.
XX
KW Tumour interacting protein; cancer; gene therapy; vector;
KW 5T4 antigen; monoclonal antibody; single chain antibody;
KW mouse; human; 5T4Sabl.
XX
OS Chimeric - Mus sp.
OS Chimeric - Homo sapiens.
OS Chimeric - synthetic.
XX
PN WO9855607-A2.
XX
PD 10-DEC-1998.
XX
```

```
PF 04-JUN-1998; 98WO-GB01627.
XX
PR 04-JUL-1997; 97GB-0014230.
PR 04-JUN-1997; 97GB-0011579.
XX 20-JUN-1997; 97GB-0013150.
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Bebbington CR, Carroll MW, Ellard FM, Kingsman SM;
PI Myers KA;
XX
DR WPI; 1999-059910/05.
XX N-PSDB; AAV80291.
PT New vector encoding a tumour interacting protein for treating cancer
PT - contains a desired nucleotide sequence and/or protein which
PT recognises tumours, and is used as a gene delivery system to treat
XX cancer
PS Example 1; Fig 1B; 82pp; English.
XX
CC This is the amino acid sequence of a single chain antibody (Sabl),
CC termed 5T4Sabl, comprising an scFv derived from murine monoclonal
CC antibody 5T4 (see AAW86002) and the human gI constant region. cDNA
CC (see AAV80291) encoding the Sabl has been inserted into vector pcNeo
CC to allow expression in mammalian cells. The trophoblast cell
CC surface antigen defined by 5T4 is expressed at high levels on the
CC cells of a wide variety of human tumours. The invention relates to
CC a vector comprising a nucleotide sequence coding for a tumour
CC interacting protein (TIP) and optionally a nucleotide sequence of
CC interest (NOI) which encodes a protein of interest (POI), the vector
CC being capable of delivering the NOI and/or POI to the tumour
CC recognised by the TIP. Delivery can be in vivo or ex vivo. The
CC vector is used to treat cancer, and may also be used as a gene
CC delivery system for introducing at least 1 gene encoding a TIP
CC (preferably a tumour binding protein) into a haematopoietic cell
XX lineage.
SQ Sequence 595 AA;
Query Match 30.4%; Score 1069; DB 20; Length 595;
Best Local Similarity 37.9%; Pred. No. 3.2e-56;
Matches 254; Conservative 20; Mismatches 60; Indels 336; Gaps 9;
QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYIMHWKSPGKLEWIGRINPNNGVTLY 60
DB 23 EVLOQSGDPLVKPGASVKISCKASGYFTGYIMHWKSPGKLEWIGRINPNNGVTLY 82
QY 61 NQKFKDKATLTVDKSTTAYMELRLSLTSDSAVYYCARSTMITNTYMDYWGQGTSTVTVSS 120
DB 83 NQKFKDKATLTVDKSTTAYMELRLSLTSDSAVYYCARSTMITNTYMDYWGQGTSTVTVS- 141
QY 121 AKTTPPSVYPLAPGSAQAQTNSMTGLCLVKGYFPEPVTVTVNSGSLSSGVHTFPAVLQSD 180
DB 142 ----- 141
QY 181 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDSGGPSEKSEINEKDLRKK 240
DB 142 -----SGG----- 144
QY 241 SELOGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTGHPWYNLLVDLGSTA 300
DB 145 ----- 144
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEEKVPIINLWIDGKQ 360
DB 145 -----GGSGGGT-----GG----- 154
QY 361 TTVPIDKVKTSKKEVTVOELDLQARHLYLHGKFLYNSDSFGCKVQVGLIVFHSSEGSTVS 420
DB 155 -----GG----- 156
QY 421 YDLFDAQGOYPTDLLRIYRDNTTISTSLISLYLYTTSIVMTQPTSLVLSAGDRVTIT 480
```


XX WO9515382-A.
PN 08-JUN-1995.
PD 29-NOV-1994; 94WO-GB02610.
PF 03-JUN-1994; 94GB-0011089.
PR 03-DEC-1993; 93GB-0024819.
XX (ZENE) ZENECA LTD.
XX
XX Blakey DC, Boot C, Copley CG, Hall SM, Paterson DS;
PI Rose MS, Wright AF;
XX WPI: 1995-215262/28.
DR
XX Antigen binding structures containing CDRs recognising the CA55.1
PT antigen - produced by hybridomas and host cells, for use in the
PT diagnosis and therapy of cancer
XX
PS Claim 3; Page 97-98; 121pp: English.
XX An antigen binding structure is based on the CDRs (given in AAR76078-
CC 84) of the heavy (AAR76085) and light (AAR76086) chains of Mab 55.1
CC (ECACC 93081901), which recognises the colorectal tumor-associated
CC antigen CA55.1. It is optionally humanized and in the form F(ab')₂,
CC F(ab)', Fab', Fv, scFv or v-min, and is produced in transgenic
CC animals or plasmids.
XX
SQ Sequence 445 AA;
Query Match 30.0%; Score 1055.5; DB 16; Length 445;
Best Local Similarity 37.5%; Pred. No. 1.5e-55;
Matches 250; Conservative 60; Mismatches 108; Indels 249; Gaps 14;
Qy 1 EVQLQQSPDLVLPKPGASVKISKASGYSFTGYMHWKQSPGKLEWIGRINPNNGVTLY 60
Db 1 QVQLQPGAEVLVLPKPGASVQLSCKASGYTFTGWIHWKQSPGKLEWIGVNPSTGRSDY 60
Qy 61 NQKFKRATLVDSSTAYWELSLTSEDSAVYCARSTMI-TNYMDYVQGTSTVTVS 119
Db 61 NEKFKRATLVDSSTAYWQLSSTSEDSAVYCARERAYGYDDAMDYVQGTSTVTVS 120
Qy 120 SAKTTPPSVYPLAPCSAQTNSMTVLGCLVKGYPEPTVTWNSGSLSSGVTTPAVLQS 179
Db 121 SAKTTPPSVYPLAPCSAQTNSMTVLGCLVKGYPEPTVTWNSGSLSSGVTTPAVLQS 180
Qy 180 DLYTLSSSVTVPSSTWPSSETVTCNVAPASSTKVDKIKVPRDSGSPSEKSEINEKDLRK 239
Db 181 DLYTLSSSVTVPSSTWPSSETVTCNVAPASSTKVDKIKVPRDCG----- 224
Qy 240 KSELQGTALGNLKIYYNYSKAITSSSEKSAQDLNTLLFKGFTGHPWYNLLVDLGST 299
Db 240 KSELQGTALGNLKIYYNYSKAITSSSEKSAQDLNTLLFKGFTGHPWYNLLVDLGST 299
Qy 225 -CKPCICTVPEVSSVFIFPPK-----PKDVLITL----- 253
Db 225 -CKPCICTVPEVSSVFIFPPK-----PKDVLITL----- 253
Qy 300 AATSEYEGSSVDLYGAYGYOCAGTGNKTAACMYGGVTLHDNRLTEKKVPIINLWIDGK 359
Db 300 AATSEYEGSSVDLYGAYGYOCAGTGNKTAACMYGGVTLHDNRLTEKKVPIINLWIDGK 359
Qy 254 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD-- 278
Db 254 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD-- 278
Qy 360 QTTVPIDVKVTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVGQGLIV 410
Db 360 QTTVPIDVKVTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVGQGLIV 410
Qy 279 -DVEVHTAQTPREEQFNSTFRSVSELPIMHQDLNKGKFCRVNSAFAFPATIEK---- 332
Db 279 -DVEVHTAQTPREEQFNSTFRSVSELPIMHQDLNKGKFCRVNSAFAFPATIEK---- 332
Qy 411 FHSSEGSTVSDLDFAQGYQPDLLRIYRDNTTISSTLSLSLYLYTTSIVMTQTPSLL 470
Db 411 FHSSEGSTVSDLDFAQGYQPDLLRIYRDNTTISSTLSLSLYLYTTSIVMTQTPSLL 470
Qy 333 -----TISK----- 337
Db 333 -----TISK----- 337
Qy 471 VSAGDRVITTCASQSVNDVANYQKPGQSPKLLISYTSRYAGVPDRFSGSGYGTDTF 530
Db 471 VSAGDRVITTCASQSVNDVANYQKPGQSPKLLISYTSRYAGVPDRFSGSGYGTDTF 530
Qy 338 ----- 337
Db 338 ----- 337
Qy 531 LTISVQAEADAAYFCQDYNPPTFGGGTKLEIKRADAPTIVSIFPPSSQLTSGGASV 590
Db 531 LTISVQAEADAAYFCQDYNPPTFGGGTKLEIKRADAPTIVSIFPPSSQLTSGGASV 590

Db 338 -----KGRPKAPQVYTIPPPKEQAKDKVSL 363
Qy 591 VCFLNNFYPKDVNVKIDGSRQGVNSWTDQDSKSTYSMSSTLTLTDEYERHNSY 650
Db 364 TCMITDFFPEDITVEMOWNGOPAEN-YKNTQPIMDT-DGSYFVYSKLVNQSNEAGNTE 421
Qy 651 TCEATHK 657
Db 422 TCSVLHE 428
RESULT 11
AAR45012
ID AAR45012 standard; protein; 230 AA.
XX AAR45012;
AC
XX 08-JUN-1994 (first entry)
DT
XX Staphylococcal enterotoxin SEE.
DE
XX Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
KW autoimmune disease; toxicity; Protein A; perfusion system.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 120 /note= "Given in the specification as J, no further
FT details given"
FT Misc-difference 121 /note= "Given in the specification as J, no further
FT details given"
FT Misc-difference 123 /note= "Given in the specification as O, no further
FT details given"
FT Misc-difference 124 /note= "Given in the specification as U, no further
FT details given"
XX WO9324136-A.
PN
XX 09-DEC-1993.
PD
XX 01-JUN-1993; 93WO-US05213.
PF
XX 01-JUN-1992; 92US-0891718.
PR
XX (STON/) STONE J L.
PA (TERM/) TERMAN D S.
XX
PI Stone JL, Terman DS;
XX
XX WPI: 1993-405418/50.
DR
XX Use of staphylococcal enterotoxin(s) and homologues - for
PT treating cancer in a patient or for the treatment of auto-immune
PT diseases
XX
PS Disclosure; Fig 1; 90pp: English.
XX
XX The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
CC which may be used in the methods of the invention for treating cancer
CC in a patient. These SEs, and homologues of them, can be used as
CC tumouricidal agents for treating cancers and autoimmune disease.
CC They exhibit tumouricidal activity and toxicity identical to that
CC observed for the Protein A perfusion system. They may be administered
CC by i.v. injection.
XX
SQ Sequence 230 AA;
Query Match 29.8%; Score 1048; DB 14; Length 230;

PT administered by IV injection, having same tumoricidal activity
PT as Staphylococcal protein A without potential toxic reactions
XX
PS Disclosure; Fig 1; 74pp; English.

XX SEE was isolated and purified from S.aureus. It can be used for
XX treating cancer, activating cytokine mediators and procoagulant
CC systems, augmenting natural killer cell cytotoxicity, etc. The
CC enterotoxin can be administered intravenously, optionally with
CC ibuprofen to attenuate toxic reaction to SEE. Synthetic
CC polyproptides having structural homology to Staphylococcal exotoxins
CC are claimed, provided the homology includes statistically
CC significant sequence homology, alignment of cysteine residues and
CC similar hydrophathy profiles.

XX Sequence 230 AA;
SQ
Query Match 29.4%; Score 1035; DB 12; Length 230;
Best Local Similarity 84.3%; Pred. No. 1.1e-54;
Matches 194; Conservative 14; Mismatches 22; Indels 0; Gaps 0;

Qy 229 SEETNEKDLRKKSELOGTALGNLQIYYNKAITSSEKSDQFLTNLLPKGFTGHPW 288
Db 1 SEETNEKDLRKKSELOGTALGNLQIYYNKAITSSEKSDQFLTNLLPKGFTGHPW 60

Qy 289 YNDLLVDLGSTAASTSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEK 348
Db 61 YNDLLVDLGSKDANKYKGVKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEX 120

Qy 349 KPNLWIDGKQTTVPIDKVKTSKKEVTQELDLQARHLYHGKFLYNSDSFGGKVORGL 408
Db 121 XVQXBKWIDGKQTTVPIDKVKTSKKEVTQELDLQARHLYHGKFLYNSDSFGGKVORGL 180

Qy 409 IVFHSSEGSTVSOLFDAQGYQPDTLRIYRDNNTISLSISLYTT 458

Db 181 IVFHSSEGSTVSOLFDAQGYQPDTLRIYRDNNTISLSISLYTT 230

RESULT 14

ID AE18377 standard; Protein; 597 AA.

XX AE18377;

XX 07-MAY-2002 (first entry)

XX Human N-terminal DAV-1 heavy chain-mature TNF-alpha fusion protein.

XX Human; bifunctional molecule; monoclonal antibody; gene therapy; cancer;
KW vascular disorder; diabetic retinopathy; restenosis; ophthalmic disorder;
KW hyperproliferative disorder; hormonal disorder; DAV-1 heavy chain;
KW cystostatic; vasotropic; ophthalmological; tumour necrosis factor-alpha;
KW TNF-alpha; fusion protein.

XX Chimeric - Homo sapiens.

XX Chimeric - Synthetic

XX Key Location/Qualifiers
FH Region 1..439
FT /note= "N-terminal portion of DAV-1 heavy chain"
FT Region 441..597
FT /note= "Human mature TNF-alpha"

XX WO200204522-A2.

XX 17-JAN-2002.

XX 09-JUL-2001; 2001WO-EP07878.

XX 10-JUL-2000; 2000US-0613017.

XX (NOVS) NOVARTIS AG.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERM GES MBH.
PA (SCRI) SCRIPPS RES INST.

XX Nemerow GR, Li E;

XX WPI; 2002-171707/22.

XX New bifunctional molecules comprising an antibody or its
XX antigen-binding portion, and a targeting agent, useful for e.g. gene
PT therapy, or for promoting adenoviral vector-mediated gene delivery to
PT cells lacking av integrins

XX Claim 15; Page 98-99; 106pp; English.

XX The present invention relates to a bifunctional molecule comprising an
XX antibody or its antigen-binding portion, and a targeting agent where the
XX antibody specifically binds to an antigen in a protein that binds to av
XX integrin, and the targeting agent specifically binds to a cell surface
XX protein that activates the phosphatidylinositol 3 (PI3K) signalling
XX pathway. The bifunctional molecules are useful for gene therapy, for
XX promoting Adenoviral (Ad) vector-mediated gene delivery to cells lacking
XX av integrins, for enhancing Ad binding and internalisation, and in gene
XX delivery of by fibreless adenovirus particles. The bifunctional molecules
XX permit targeting of viral and bacterial vectors to cells that express,
XX targeted receptors. Diseases that can be targeted include cancers,
XX vascular disorders, diabetic retinopathies, restenosis, ophthalmic
XX disorders, hyperproliferative disorders, and hormonal disorders.
XX The present sequence is human N-terminal DAV-1 heavy chain-mature tumour
XX necrosis factor-alpha (TNF-alpha) fusion protein which is used in the
XX invention.

XX Sequence 597 AA;

XX Query Match 29.2%; Score 1028; DB 23; Length 597;

XX Best Local Similarity 37.2%; Pred. No. 9.4e-54;

XX Matches 253; Conservative 53; Mismatches 115; Indels 260; Gaps 15;

Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSFTGYMHVWKQSPGKGLEWIGRINPNNGVTLY 60

Db 20 EVQLQQSGPELVKPGASVKISKASGYTFTDYNMHVWKQSHGKSLIEWIGYIPYKGGTGY 79

Qy 61 NQPKFKDRLTLVDKSSVTAYMELRSLTSEDSAVYYCARSTMTITNYVDYWGQGSVTYSS 120

Db 80 NQPKFKDRLTLTSSNTAYMELRSLTSDASAVYYCARG-----IAYWGQGLVTYSA 132

Qy 121 AKTTPPSVYPLAPGSAATNSMVTLGCLVKGYFPEPVTVTWNSGSLSGVHTFPVAVLQSD 180

Db 133 AKTTPPSVYPLAPGSAATNSMVTLGCLVKGYFPEPVTVTWNSGSLSGVHTFPVAVLQSD 192

Qy 181 LYTLSSTVTPSSWPSETVTCNVAHPASSTKVDKKIIPRDSGGPSEKSEEINEKDLRKK 240

Db 193 LYTLSSTVTPSSWPSETVTCNVAHPASSTKVDKKIIPRDCG----- 235

Qy 241 SELQGTALGNLKQIYYNKAITSSEKSDQFLTNLLFKGFTGHTGHPWYNDLLVDLGSTA 300

Db 236 CKPCICTVPEVSSVFIFPPK-----PKDVLATLT----- 264

Qy 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPTINLWIDCKQ 360

Db 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWEVD--- 289

Qy 361 TTVPIDKVKTSKKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVQRLIVF 411

Db 290 -DVEVHTAQTPREQPFNSERSVSELPFIHQDWLNGREFKCRVNSAAPPAPIEK----- 343

Qy 412 HSEGSTVSVDLFDAGQGYPDTLRIYRDNNTISLSISLYLTTISIVMTQTPTSLV 471

Db 344 -----TISKT----- 348

Qy 472 SAGDRVTITCKASQSVNDVAWYQKPGOSPKLLISYTSRYAGVYVDRFSGSGYGTDFTL 531

Db 349 ----- 348

QY 532 TISSVQAEADAAVYFCQDDYNSPPTFGGKLEIKRADAAPTVSIFFPPSEQLTSGGASVY 591
Db 349 -----KGRKAPQVYTIPTPPKQEQMAKDKVSLT 375
QY 592 CFLNFFPKDINVKIKDGSERQGNLSWTDQDSDKSTYSMSSTLTITKDYERHNSYT 651
Db 376 CMITDFFPDITVEQWNGQPAEN-YKNTQPIMDT-DGSFYVYSKLNVOKSNWEAGNTFI 433
QY 652 CEATHKTSPIVKSFNRNES 672
Db 434 CSVLHE-----FVRSSSRTPS 449
RESULT 15
ID AAW83041
XX AAW83041 standard; Protein; 464 AA.
AC AAW83041;
XX
DT 15-MAR-1999 (first entry)
DE Anti-Fas MAb HFE7A heavy chain.
XX
KW HFE7A; monoclonal antibody; mouse; Fas; humanised antibody;
KW apoptosis; HFE7A; autoimmune disease; Hashimoto's disease;
KW systemic lupus erythematosus; graft versus host disease;
KW Sjogren syndrome; pernicious anaemia; Addison's disease;
KW scleroderma; Goodpasture syndrome; Crohn's disease; sterility;
KW rheumatoid arthritis; autoimmune haemolytic anaemia;
KW myasthenia gravis; multiple sclerosis; Basedow's disease;
KW thrombopenia purpura; insulin-dependent diabetes; allergy;
KW atopy; arteriosclerosis; myocarditis; cardiomyopathy;
KW glomerular nephritis; hypoplastic anaemia; hepatitis; AIDS;
KW transplant rejection; therapy; complementarity determining region;
CW CDR.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Peptide 1..19
FT /label= Sig_peptide
FT Protein 20..464
FT /label= Mat_protein
FT Region 20..140
FT /label= Variable
FT Region 141..464
FT /label= Constant
FT Region 50..54
FT /label= CDR_H1
FT /note= "claim 9"
FT Region 69..84
FT /label= CDR_H2
FT /note= "claim 9"
FT Region 118..128
FT /label= CDR_H3
FT /note= "claim 9"
XX AU9859701-A.
XX
XX 08-OCT-1998.
XX
XX 30-MAR-1998; 98AU-0059701.
XX
XX 08-OCT-1997; 97JP-0276064.
PR 01-APR-1997; 97JP-0082953.
PR 25-JUN-1997; 97JP-0169088.
XX
XX (SANY) SANKYO CO LTD.
XX
XX Akio S, Hideyuki H, Hiroko Y, Jun O, Kimihisa I;
PI Masahiko O, Nobufusa S, Shin Y, Tohru T;
XX
XX WPI: 1998-543440/47.

DR N-PSDB: AAV71029.
XX New antibodies and proteins bind conserved epitope of Fas antigen -
PT used to evaluate drugs in animal models and to treat Fas-associated
PT diseases e.g. autoimmune disease, allergy, atopy, arteriosclerosis,
PT myocarditis, hepatitis and AIDS
XX
PS Reference Example 4; Page 187-188; 292pp: English.
XX
XX This is the amino acid of the heavy chain of murine anti-human Fas
CC monoclonal antibody HFE7A. cDNA (see AAV70129) encoding the heavy
CC chain was obtained from HFE7A-secreting hybridoma (FERM BP-5828)
CC RNA by RT-PCR (see AAV70125-26). The invention provides humanised
CC HFE7A antibodies (see AAW83031-37) produced by CDR grafting. These
CC antibodies are capable of inducing apoptosis in abnormal cells
CC expressing Fas, and of inhibiting Fas-induced apoptosis in normal
CC cells. They are used to evaluate, in animal models, treatments of
CC diseases that involve Fas/Fas ligand interactions, and also to treat
CC such diseases, including autoimmune disease (e.g. systemic lupus
CC erythematosus, Hashimoto's disease, graft versus host disease,
CC Sjogren syndrome, pernicious anaemia, Addison's disease,
CC scleroderma, Goodpasture syndrome, Crohn's disease, rheumatoid
CC arthritis, autoimmune haemolytic anaemia, sterility, myasthenia
CC gravis, multiple sclerosis, Basedow's disease, thrombopenia purpura
CC and insulin-dependent diabetes), allergies, atopy, arteriosclerosis,
CC myocarditis, cardiomyopathy, glomerular nephritis, hypoplastic
CC anaemia, hepatitis, AIDS and transplant rejection (all claimed).
XX
SQ Sequence 464 AA;
Query Match 29.1%; Score 1025.5; DB 19; Length 464;
Best Local Similarity 36.2%; Pred. No. 9.8e-54;
Matches 243; Conservative 64; Mismatches 106; Indels 259; Gaps 14;
QY 1 EVQLQSGPDIVKPGASVKISKASGYSTFTGYMHVWVQSPCKGLEWIGRINPNNGVTLY 60
Db 20 QVQLQPGAEIVKPGASVKLSCKASGYFTSYMMQWVKRPGQGLEWIGIDPSDSTNY 79
QY 61 NQKFKDKATLTVDKSTTAYMELSLTSDSAVYYCARS-TWITNYVMYWGQTSVTYS 119
Db 80 NQKFKGKATLTVDTSSTAYMQLSSTSDSAVYYCARNRDYSNNWYFDVWGTTVTYS 139
QY 120 SAKTTPPSVYPLAPGSAAGTNSMVTLCGLVKGYFPEPVTVTWNSGSLSGVHTFPAVLQS 179
Db 140 SAKTTPPSVYPLAPGSAAGTNSMVTLCGLVKGYFPEPVTVTWNSGSLSGVHTFPAVLQS 199
QY 180 DLYTLSSSVTPSPSTWPSSETVTCNVAHPASSTKVDKIVPRDSGGPSEKSEINEKDLRK 239
Db 200 DLYTLSSSVTPSPSTWPSQTVCNVAHPASSTKVDKIVPRDCG----- 243
QY 240 KSELQGTALGNLKOIYYNYSKAITSEKSAQDQFLNTLLFKGFTGHPWYNLLVDLGST 299
Db 244 -CKPCICTVPEVSSVFIFPPK-----PKDVLITL----- 272
QY 300 AATSEYEGSSVDLYGAYYGYQCAGGTPNKACMYGGVTLHDNNRLTEKKVPINLWIDGK 359
Db 273 -----TP-KVTCVWVDIS-----KDDPEVQFSNEVD-- 297
QY 360 QTTVPIDKVKTSKKEV-----TVQELDLQARHILHCK-FGL-YNSDSFGCKVQGLIV 410
Db 298 --DEVHTAQTPREEQFNSTERSVSELPIMHQNLNGKEKFCRVNSAFAPIEK----- 351
QY 411 FHSSEGSTVSYDLFDAQGYPDTLRLIYRDNTTISSTLSLSLYLVTTYSIVMTQPTSL 470
Db 352 -----TISKT----- 356
QY 471 VSAGDRVITITKASQSVSNDAVYQKQPSKLLISYTSRYAGVDPDRFSGSGYGTDF 530
Db 357 ----- 356
QY 531 LTISVQAEADAAVYFCQDDYNSPPTFGGKLEIKRADAAPTVSIFFPPSEQLTSGGASV 590
Db 357 -----KGRKAPQVYTIPTPPKQEQMAKDKVSL 382

```

QY 591 VCLNNFYPKDINVKWKIDGSRQN-----GVLNSWTQDQSKDSTYSMSSTLTLLTKDEYE 645
  | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 383 TCMTIDFFPEDITVEQWNGQPAENYKNTQPIMNT-----NGSYFYYSKLVNOKSNWE 435

QY 646 RHNSYTCEATHK 657
  | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 436 AGNTFTCSVLHE 447

RESULT 16
AAB14747
ID AAB14747 standard; Protein: 464 AA.
AC AAB14747;
DT 24-NOV-2000 (first entry)
DE Mouse anti-Fas antibody HFE7A heavy chain.
XX
KW Anti-Fas antibody; monoclonal antibody HFE7A; FERM-BP-5828;
KW murine; complementarity determining region; CDR; human Fas;
KW Fas ligand; apoptosis modulator; programmed cell death;
KW autoimmune disease; allergy; atopy; arteriosclerosis; myocarditis;
KW cardiomyopathy; glomerulonephritis; aplastic anaemia; pancytopenia;
KW hepatitis; AIDS; graft rejection; heavy chain.
XX
OS Mus musculus.
XX
PN JP2000169393-A.
XX
PD 20-JUN-2000.
XX
PF 30-SEP-1999; 95JP-0278301.
XX
PR 30-SEP-1998; 98JP-0276883.
XX
PA (SANY ) SANKYO CO LTD.
XX
WPI: 2000-485645/43.
DR N-PSDB; AAA72108.
XX
PT Preventive or treating agent for the diseases caused by an abnormality
PT in the Fas/Fas ligand system e.g. autoimmune diseases, contains
PT anti-Fas antibody.
XX
PS Example 4; Page 67-68; 139pp; Japanese.
XX
CC The invention relates to compositions for the prevention or treatment
CC or diseases caused by an abnormality in the Fas/Fas ligand system
CC containing an anti-Fas antibody as the active component. The anti-Fas
CC antibody is either the murine anti-human Fas monoclonal antibody HFE7A,
CC or a humanised version of HFE7A containing identical CDRs
CC (complementarity determining regions) to antibody HFE7A. Via its
CC interaction with Fas, the antibody of the invention acts as a modulator
CC of apoptosis. The compositions of the invention may therefore be used in
CC the treatment or prevention of conditions such as autoimmune diseases,
CC allergy, atopy, arteriosclerosis, myocarditis, cardiomyopathy,
CC glomerulonephritis, aplastic anaemia (panmyelophthisis), hepatitis, AIDS
CC and organ graft rejection. The present sequence represents the
CC heavy chain of the murine anti-human Fas monoclonal antibody HFE7A,
CC which is produced by hybridoma HFE7A (FERM-BP-5828).
XX
SQ Sequence 464 AA;

Query Match 29.1%; Score 1025.5; DB 21; Length 464;
Best Local Similarity 36.2%; Pred. NO. 9.8e-54;
Matches 243; Conservative 64; Mismatches 106; Indels 259; Gaps 14;

QY 1 EVLOQSGPDLVKPCAGKISKASGYSTGYGMHWKQSPCKGLEWIGRINPNNGVTLY 60
  : : : : : | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 20 QVLOQPGAEVLKPGASVKLSCKASGYFTTSYMQMVKQRPQGGLIEWIGIDPSDSYTN 79

```

```

QY 61 NQKFKDKATLTVDKSSTTAYMELRSLTSEDSAVYYCARS-TMITNYVMYMQSGTSTVTS 119
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 80 NQKFKGKATLVDTSSTAYMQLSSLTSEDSAVYYCARNDYSNWYFDVWGTGTTVTS 139

QY 120 SAKTTPPSVYPLAPGSAAGTNSMVTGLCLVKGYPPEPVTVTWNSGSLSSGVHFFPAVLQS 179
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 140 SAKTTPPSVYPLAPGSAAGTNSMVTGLCLVKGYPPEPVTVTWNSGSLSSGVHFFPAVLQS 199

QY 180 DLYTLSSSVTPSPSTPSETVTCNVAHPASSPKVDKKLVPRDSGGPSEKSEINEKDLRK 239
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 200 DLYTLSSSVTPSPSTPSETVTCNVAHPASSPKVDKKLVPRDCG----- 243

QY 240 KSELQGTALGNLKIYYNKAITSSEKSADQFLNTLLFKGFFTGHPWYNLLVDLGST 299
  | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 244 -CKPCICTVPEVSSVFIFPK-----PKDVLITL----- 272

QY 300 AATSEYEGSSVDLYGAYYGOCAGGTPNKTACMYGGVTLHDNRLTEBKKVPINLWIDGK 359
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 273 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD-- 297

QY 360 QTTVPIDKVKTSKKEV-----TYQELDLQARHYLHGK-FGL-YNSDSFGKVKVQGLIV 410
  | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 298 --DVEHTAQTPREEQNFSTRSVSELPIMHQNWLNGKEFKRCRVNSAAFPAPIEK----- 351

QY 411 FHSSEGSTSVSYDLFDAQGYPTDLLRIYRDNTTISSTLSLSISLYLTTSIVMTQPTSL 470
  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 352 -----TISKT----- 356

QY 471 VSAGDRVITTCASOSVNDVAVYOOKPQSPKLLISYTSRYAGVPRFSGSGYGTDT 530
  | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 357 ----- 356

QY 531 LTISSVQAEADAAVYFCQDYNSPPTFGGKLEIKRADAAPTVSIFPPSSSEQLTSGGASV 590
  | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 357 -----KGRPKAPQVYTPPPKEQMAKDKVSL 382

QY 591 VCLNNFYPKDINVKWKIDGSRQN-----GVLNSWTQDQSKDSTYSMSSTLTLLTKDEYE 645
  | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 383 TCMTIDFFPEDITVEQWNGQPAENYKNTQPIMNT-----NGSYFYYSKLVNOKSNWE 435

QY 646 RHNSYTCEATHK 657
  | : : : : | : : : : | : : : : | : : : : | : : : : | : : : : |
Db 436 AGNTFTCSVLHE 447

RESULT 17
AAW90897
ID AAW90897 standard; Protein: 464 AA.
XX
AC AAW90897;
XX
DT 08-AUG-2000 (first entry)
XX
DE Murine anti-Fas antibody HFE7A heavy chain protein.
XX
KW Fas; antibody; murine; anti-inflammatory; anti-anemic; antidiabetic;
KW anti-allergic; anti-arthritis; antiviral; immunomodulatory; cardiac;
KW dermatological; immunosuppressive; thyromimetic; antirheumatic; anti-
KW nephrotropic; antifertility; neuroprotective; antiarteriosclerotic;
KW hepatotropic; humanized; apoptosis; systemic lupus erythematosus; HFE7A;
KW Hashimoto disease; rheumatoid arthritis; graft versus host disease;
KW Sjogren's syndrome; anemia; Addison's disease; scleroderma; sterility;
KW Goodpasture syndrome; Crohn's disease; sterility; myasthenia gravis;
KW multiple sclerosis; Basedow's disease; thrombopenia purpura; allergy;
KW insulin dependent diabetes mellitus; arteriosclerosis; myocarditis;
KW cardiomyopathy; glomerulonephritis; hepatitis; transplant rejection.
XX
OS Mus musculus.
XX
PN EP990663-A2.
XX
PD 05-APR-2000.
XX

```



```

591 VCFLNNFYPKDINVKWKIDGSRQN-----GVLSNWTQDQSKDSTYSMSSTLTLTKEVE 64
| : : : : | : : : : | : : : : | : : : : | : : : : |
357 -----KGRPAQPVIIITLPPFPKQEQMAKDRAVSL 36

```



```

FH Key      Location/Qualifiers
FT Region   230..242
          /note= "Hinge region"
XX
PN WO200204522-A2.
PD
XX 17-JAN-2002.
XX
XX 09-JUL-2001; 2001WO-EP07878.
XX
XX 10-JUL-2000; 2000US-0613017.
XX
XX (NOVS ) NOVARTIS AG.
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PA (SCRI ) SCRIPPS RES INST.
XX
XX Nemerow GR, Li E;
XX
XX WPI; 2002-171707/22.
DR N-PSDB; AAD29308.
XX
XX New bifunctional molecules comprising an antibody or its
PT antigen-binding portion, and a targeting agent, useful for e.g. gene
PT therapy, or for promoting Adenoviral vector-mediated gene delivery to
PT cells lacking av integrins
XX
XX Claim 10; Page 91-92; 106pp; English.
XX
XX The present invention relates to a bifunctional molecule comprising an
CC antibody or its antigen-binding portion, and a targeting agent where the
CC antibody specifically binds to an antigen in a protein that binds to av
CC integrin, and the targeting agent specifically binds to a cell surface
CC protein that activates the phosphatidylinositol 3 (PI3K) signalling
CC pathway. The bifunctional molecules are useful for gene therapy, for
CC promoting Adenoviral (Ad) vector-mediated gene delivery to cells lacking
CC av integrins, for enhancing Ad binding and internalisation, and in gene
CC delivery of by fibreless adenovirus particles. The bifunctional molecules
CC permit targeting of viral and bacterial vectors to cells that express
CC targeted receptors. Diseases that can be targeted include cancers,
CC vascular disorders, diabetic retinopathies, restenosis, ophthalmic
CC disorders, hyperproliferative disorders, and hormonal disorders.
CC The present sequence is human penton base monoclonal antibody, DAV-1
CC heavy chain.
XX
XX Sequence 456 AA;
SQ
Query Match      29.1%; Score 1023.5; DB 23; Length 456;
Best Local Similarity 37.4%; Pred. No. 1.3e-53;
Matches 249; Conservative 51; Mismatches 111; Indels 255; Gaps 14;
QY 1 EVLOQSGPDLVKPGASVKASGYSFTGYMHWKSPGKLEWIGRIINPNNGVTLY 60
DB 20 EVLOQSGPDLVKPGASVKASGYSFTGYMHWKSPGKLEWIGRIINPNNGVTLY 79
QY 61 NQKFKDKATLTVDKSTTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGQGTSTVSS 120
DB 80 NQKFKSKATLTDTSSNTAYMELRSLTSDSAVYYCARG-----IAYWGQGTLTVSA 132
QY 121 AKTTPSPVYPLAPGSAQAQNSWVTLGCLIVKGYFPEPVTVWNSGSLSSGVHFTFAVLQSD 180
DB 133 AKTTPSPVYPLAPGSAQAQNSWVTLGCLIVKGYFPEPVTVWNSGSLSSGVHFTFAVLQSD 192
QY 181 LYTLSSTVTPSTWSPSEVTNCVAHPASSTKVDKIVPRDSDGSPSEKSEINEKDLRKK 240
DB 193 LYTLSSTVTPSTWSPSEVTNCVAHPASSTKVDKIVPRDCG----- 235
QY 241 SELQGTALGNLKOIYYNYSKAITSSSEKSAQDLFTNTLFFKGFETGHPWYNLLDLVGSTA 300
DB 236 CKPCICTVPEVSSVFPPK-----PKDVLITL----- 264
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
DB 265 -----TP-KVTCVVVDIS-----KDDPEVQFSEFVD--- 289
```

```

QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGKVGQGLIVF 411
DB 290 -DVEVHTAQTPREECFNSTFRSVSELPIMHODWLNKGEKCRVNSAAPAEIK----- 343
QY 412 HSSEGSTVSYDLFDAGQGYPDTLRLRIYRDNTTISSTLSLSISLYLYTTSIVMTQTPTSLLV 471
DB 344 -----TISKT----- 348
QY 472 SAGDRVTITCKASQSVSNVDVANYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDFTL 531
DB 349 ----- 348
QY 532 TISSVQAEDAAYVFCQDYNSPPTGGGPKLEIKRADAAPTYSIPPPSEQLTSGGASVV 591
DB 349 -----KGRPKAPQVYTTIPPPKEQMAKKVSLT 375
QY 592 CFLNMFYPKDINVKWKIDGSRQNGVLNSWTQDQSKDSTYSMSSTLTLTQDEYERHNSYT 651
DB 376 CMITDFPEPEDITVEQWNGQPAEN-YKNTPIMDT-DGSYFYVSKLNVQKSNWEAGNTFI 433
QY 652 CEATHK 657
DB 434 CSVLHE 439
RESULT 22
AAE18379
ID AAE18379 standard; Protein; 493 AA.
XX
AC AAE18379;
DT 07-MAY-2002 (first entry)
XX
DE Human N-terminal DAV-1 heavy chain-mature EGF fusion protein.
XX
KW Human; bifunctional molecule; monoclonal antibody; gene therapy; cancer;
KW vascular disorder; diabetic retinopathy; restenosis; ophthalmic disorder;
KW hyperproliferative disorder; hormonal disorder; DAV-1 heavy chain;
KW cytostatic; vasotropic; ophthalmological; epidermal growth factor;
KW EGF; fusion protein.
XX
OS Chimeric - Homo sapiens.
OS Chimeric - Synthetic
XX
FH Key      Location/Qualifiers
FT Region   1..439
          /note= "N-terminal portion of DAV-1 heavy chain"
FT Region   441..493
          /note= "Human mature EGF"
XX
XX WO200204522-A2.
XX
XX 17-JAN-2002.
XX
XX 09-JUL-2001; 2001WO-EP07878.
XX
XX 10-JUL-2000; 2000US-0613017.
XX
XX (NOVS ) NOVARTIS AG.
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PA (SCRI ) SCRIPPS RES INST.
XX
XX Nemerow GR, Li E;
XX
XX WPI; 2002-171707/22.
XX
XX New bifunctional molecules comprising an antibody or its
PT antigen-binding portion, and a targeting agent, useful for e.g. gene
PT therapy, or for promoting Adenoviral vector-mediated gene delivery to
PT cells lacking av integrins
XX
XX Claim 15; Page 101-102; 106pp; English.
```

RESULT_23	
AAE18378	
ID	AAE18378 standard; Protein; 510 AA.
XX	
XX	AAE18378;
XX	
XX	
XX	07-MAY-2002 (first entry)
XX	
XX	
DE	Human N-terminal DAV-1 heavy chain-mature IGF-1 fusion protein.
XX	
KW	Human; bifunctional molecule; monoclonal antibody; gene therapy; cancer;
KW	vascular disorder; diabetic retinopathy; restenosis; ophthalmic disorder;
KW	hyperproliferative disorder; hormonal disorder; DAV-1 heavy chain;
KW	cytostatic; vasotropic; ophthalmological; insulin growth factor-1;
KW	IGF-1; fusion protein.
XX	
OS	Chimeric - Homo sapiens.
OS	Chimeric - Synthetic

FT	Region	Location/Qualifiers
FT	Region	1..439
FT		/note= "N-terminal portion of DAV-1 heavy chain"
FT	Region	441..510

XX
PN
W0200204522-A2.

PD 17-JAN-2002.
XX
XX
PF 09-JUL-2001; 2001WO-EP07878.

10-JUL-2000; 2000US-0613017.

PA (NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.
PA (SCRI) SCRIPPS RES INST.
XX
XX

XX
DR
WPI; 2002-171707/22.
vv

PT New bifunctional molecules comprising an antibody or its
PT antigen-binding portion, and a targeting agent, useful for e.g. gene
PT therapy, or for promoting Adenoviral vector-mediated gene delivery to

XX
PS
v v

CC The present invention relates to a bifunctional molecule comprising an
CC antibody or its antigen-binding portion, and a targeting agent where the
CC antibody specifically binds to an antigen in a protein that binds to a
CC antigen.

CC inorganic, and the targeting agent specifically binds to a cell surface
CC protein that activates the phosphatidylinositol 3 (PI3K) signalling
CC pathway. The bifunctional molecules are useful for gene therapy, for
CC example, adenoviral-mediated gene delivery.

CC av integrins, for enhancing Ad binding and internalisation, and in gene
CC delivery of by fibreless adenovirus particles. The bifunctional molecules
CC permit targeting of viral and bacterial vectors to cells that express
CC

CC targeted receptors. Diseases that can be targeted include cancers, CC vascular disorders, diabetic retinopathies, restenosis, ophthalmic CC disorders, hyperproliferative disorders, and hormonal disorders. CC antibodies, including antibodies to receptors, are used in the treatment of

CC	growth factor-1 (IGF-1) fusion protein which is used in the invention.
XX	
XX	
Sequence	510 AA;

Query Match	29.1%	Score	1023.5;	DB	23;	Length	510;
Best Local Similarity	37.4%	Pred. No.	1.5e-53;				

QY 1 EVQLQQSGPDLVPGASVKISCKASGYSETGYMHVWVKQSPGKGLEWIGRINPNNGVTLY 60

20 EVQDQSGPELVKPGASVKISCKASGIIIPIDINHHWVKQSGKSEWLGIIIPINQGGIGI 79

```
QY 61 NQFKDKATLTVDKSTTAYMELRLTSDSAVYVCARSTMTNVMYDYGQGTSTVSS 120
PI ||||| ||||| || ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 80 NQFKSKATLTDTSSNTAYMELRLTSDSAVYVCARG-----IAYWGGTLTVTSA 132
DR WPI: 2002-171707/22.
DR N-PSDB: AAE18372.
XX
PT New bifunctional molecules comprising an antibody or its
PT antigen-binding portion, and a targeting agent, useful for e.g. gene
PT therapy, or for promoting Adenoviral vector-mediated gene delivery to
PT cells lacking av integrins
XX
PS Claim 10; Page 96; 106pp; English.
XX
CC The present invention relates to a bifunctional molecule comprising an
CC antibody or its antigen-binding portion, and a targeting agent where the
CC antibody specifically binds to an antigen in a protein that binds to av
CC integrin, and the targeting agent specifically binds to a cell surface
CC protein that activates the phosphatidylinositol 3 (PI3K) signalling
CC pathway. The bifunctional molecules are useful for gene therapy, for
CC promoting Adenoviral (Ad) vector-mediated gene delivery to cells lacking
CC av integrins, for enhancing Ad binding and internalisation, and in gene
CC delivery of by fibreless adenovirus particles. The bifunctional molecules
CC permit targeting of viral and bacterial vectors to cells that express
CC vascular disorders, diabetic retinopathies, restenosis, ophthalmic
CC disorders, hyperproliferative disorders, and hormonal disorders.
CC The present sequence is human penton base monoclonal antibody, DAV-1
CC heavy chain fragment.
XX
SQ Sequence 438 AA;
Query Match 29.0%; Score 1022.5; DB 23; Length 438;
Best Local Similarity 37.4%; Pred. No. 1.4e-53;
Matches 249; Conservative 50; Mismatches 111; Indels 255; Gaps 14;
QY 1 EVQLQQSGDPLVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGRINPNNGVTLY 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 20 EVQLQQSGPELVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGYIYKGGTGY 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 NQFKDKATLTVDKSTTAYMELRLTSDSAVYVCARSTMTNVMYDYGQGTSTVSS 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 80 NQFKSKATLTDTSSNTAYMELRLTSDSAVYVCARG-----IAYWGGTLTVTSA 132
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 AKTTPPSVYPLAPGSAAGTNSMTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 133 AKTTPPSVYPLAPGSAAGTNSMTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 192
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVKKIVPRDSGGPSEKSEINEKDLRKK 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 193 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVKKIVPRDCG----- 235
QY 241 SELQGTALGNLKOIYYNYSKAITSEKSDAQFLTNTLLFKGFTGHPWYNDLLVLGSTA 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 236 CKPCICTVEVSSVFIFPPK-----PKDVLITL----- 264
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTGNKTCACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWEVD----- 289
QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVQGLIYF 411
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHODWLNKKEKCRVNSAAPPAPIEK----- 343
QY 412 HSEGSTVSVDLFDAGQGYPDTLRLIYRDNTTISSTLSISLYLTTSIVMTQTPTSLIV 471
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 344 -----TISK----- 348
QY 472 SAGDRVTITCKASQSVNSDVAMVYQKPGSQKLLISYTSRYAGVPDRFSGSGYGTDFTL 531
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 349 ----- 348
QY 532 TISSVQAEDAAYVFCQDYNSPPTFGGKTKLEIKRADAAPTYSIPPSSEQLTSGGASVY 591
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 592 CFLNNFPRDINVKWKIDGSRQNGVLNSWTDQDSKDSYMSSTLTLPKDEYERHNSYT 651
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 376 CMITDFPEPDIWQWNGQPAEN-YKNTQPIMDT-DGSYFYYSKLVNQSWEAGNTFI 433
QY 652 CEATHK 657
Db |||||
QY 434 CSVLHE 439
RESULT 24
ID AAE18372 standard; Protein: 438 AA.
AC AAE18372;
DT 07-MAY-2002 (first entry)
XX Human penton base monoclonal antibody, DAV-1 heavy chain fragment.
XX Human; bifunctional molecule; monoclonal antibody; gene therapy; cancer;
KW vascular disorder; diabetic retinopathy; restenosis; ophthalmic disorder;
KW hyperproliferative disorder; hormonal disorder; DAV-1 heavy chain;
KW cytostatic; vasotropic; ophthalmological.
XX
OS Homo sapiens.
XX
PN WO200204522-A2.
PD 17-JAN-2002.
XX
PF 09-JUL-2001; 2001WO-EP07878.
XX
PR 10-JUL-2000; 2000US-0613017.
XX
PA (NOVS ) NOVARTIS AG.
PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
```

```
PA (Scri ) SCRIPPS RES INST.
XX Nemerow GR, Li E;
XX WPI: 2002-171707/22.
DR N-PSDB: AAE18372.
XX
PT New bifunctional molecules comprising an antibody or its
PT antigen-binding portion, and a targeting agent, useful for e.g. gene
PT therapy, or for promoting Adenoviral vector-mediated gene delivery to
PT cells lacking av integrins
XX
PS Claim 10; Page 96; 106pp; English.
XX
CC The present invention relates to a bifunctional molecule comprising an
CC antibody or its antigen-binding portion, and a targeting agent where the
CC antibody specifically binds to an antigen in a protein that binds to av
CC integrin, and the targeting agent specifically binds to a cell surface
CC protein that activates the phosphatidylinositol 3 (PI3K) signalling
CC pathway. The bifunctional molecules are useful for gene therapy, for
CC promoting Adenoviral (Ad) vector-mediated gene delivery to cells lacking
CC av integrins, for enhancing Ad binding and internalisation, and in gene
CC delivery of by fibreless adenovirus particles. The bifunctional molecules
CC permit targeting of viral and bacterial vectors to cells that express
CC vascular disorders, diabetic retinopathies, restenosis, ophthalmic
CC disorders, hyperproliferative disorders, and hormonal disorders.
CC The present sequence is human penton base monoclonal antibody, DAV-1
CC heavy chain fragment.
XX
SQ Sequence 438 AA;
Query Match 29.0%; Score 1022.5; DB 23; Length 438;
Best Local Similarity 37.4%; Pred. No. 1.4e-53;
Matches 249; Conservative 50; Mismatches 111; Indels 255; Gaps 14;
QY 1 EVQLQQSGDPLVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGRINPNNGVTLY 60
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 20 EVQLQQSGPELVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGYIYKGGTGY 79
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 61 NQFKDKATLTVDKSTTAYMELRLTSDSAVYVCARSTMTNVMYDYGQGTSTVSS 120
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 80 NQFKSKATLTDTSSNTAYMELRLTSDSAVYVCARG-----IAYWGGTLTVTSA 132
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 121 AKTTPPSVYPLAPGSAAGTNSMTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 133 AKTTPPSVYPLAPGSAAGTNSMTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 192
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 181 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVKKIVPRDSGGPSEKSEINEKDLRKK 240
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 193 LYTLSSTVTPSPSTWPSSETVTCNVAHPASSTKVKKIVPRDCG----- 235
QY 241 SELQGTALGNLKOIYYNYSKAITSEKSDAQFLTNTLLFKGFTGHPWYNDLLVLGSTA 300
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 236 CKPCICTVEVSSVFIFPPK-----PKDVLITL----- 264
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTGNKTCACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWEVD----- 289
QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVQGLIYF 411
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHODWLNKKEKCRVNSAAPPAPIEK----- 343
QY 412 HSEGSTVSVDLFDAGQGYPDTLRLIYRDNTTISSTLSISLYLTTSIVMTQTPTSLIV 471
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 344 -----TISK----- 348
QY 472 SAGDRVTITCKASQSVNSDVAMVYQKPGSQKLLISYTSRYAGVPDRFSGSGYGTDFTL 531
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 349 ----- 348
QY 532 TISSVQAEDAAYVFCQDYNSPPTFGGKTKLEIKRADAAPTYSIPPSSEQLTSGGASVY 591
```


QY	529	FTLTISVQAEADAAVFCQDYNSPTFGGTTKLEIKRADAAPTVSIFPPSSQLTSGGA	588
Db	63	FTTISTVQAEADLAIVFCQDYNSPTFGGTTKLEIKRADAAPTVSIFPPSSQLTSGGA	1248
QY	589	SVVCFNLNFPKIDINKWKIDGSEKQNGVLNSWTDDSKDSTYSMSSTLTLTCKDEYERHN	648
Db	123	SVVCFNLNFPKIDINKWKIDGSEKQNGVLNSWTDDSKDSTYSMSSTLTLTCKDEYERHN	1824
QY	649	SYTCEATHKSTSTSPIVKSFNRE 671	
Db	183	SYTCEATHKSTSTSPIVKSFNRE 205	
RESULT 26			
AAR66758			
XX	ID	AAR66758 standard; Protein; 465 AA.	
AC	AC	AAR66758;	
XX	DT	01-SEP-1995 (first entry)	
XX	DE	Anti-tobacco mosaic virus monoclonal Ab heavy chain.	
XX	XX		
KW	XX	Tobacco mosaic virus; TMV; monoclonal antibody;	
KW	KW	heavy chain; virus-resistant plants; biofarming.	
XX	OS	Synthetic.	
XX	Key	Location/Qualifiers	
FX	Peptide	1..19	
FT	Peptide	/label= leader	
FT	Peptide	20..465	
FT	Peptide	/label= mat_peptide	
FT	Domain	20..128	
FT	Domain	/note= "variable heavy domain"	
FT	Domain	129..141	
FT	Domain	/note= "J heavy 4 domain"	
FT	Domain	142..465	
FT	Domain	/note= "constant heavy domain"	
XX	XX		
PN	XX	JP06319396-A.	
XX	XX		
PD	XX	22-NOV-1994.	
XX	XX		
PF	XX	07-MAY-1993; 93JP-0131208.	
XX	XX		
PR	XX	07-MAY-1993; 93JP-0131208.	
XX	XX		
PA	XX	(NISB) JAPAN TOBACCO INC.	
PA	XX	(KURS) KURARAY CO LTD.	
XX	XX		
DR	XX	WPI: 1995-040220/06.	
DR	XX	N-PSDB: AAQ79930.	
XX	XX		
PT	XX	Transformed plant producing animal-derived anti-virus antibody -	
PT	XX	esp. tobacco plants producing anti-tobacco mosaic virus	
PT	XX	monoclonal antibody	
XX	XX		
PS	XX	Example 2; Pages 14-15; 26pp; Japanese.	
XX	XX		
CC	XX	AAQ79929 and AAQ79930 encode AAR66757 and AAR66758, the light and heavy	
CC	XX	chains of an animal derived anti-tobacco mosaic virus (TMV)	
CC	XX	monoclonal antibody. The cDNAs were incorporated into a Ti	
CC	XX	plasmid vector, which was incorporated into A. tumefaciens.	
CC	XX	The resultant plant expression vector was used to transform	
CC	XX	tobacco plants, making them TMV resistant, the plants could	
CC	XX	also be biofarmed for the prodn. of anti-virus antibodies.	
XX	XX		
SQ	Sequence	465 AA;	
Query Match 28.9%; Score 1018; DB 16; Length 465;			
Best Local Similarity 36.3%; Pred. No. 2.8e-53;			
Matches 244; Conservative 60; Mismatches 109; Indels 260; Gaps			


```
CC arthritis, or conditions associated with hypoxia, hypoglycemia or
CC ischemia, or to deliver antibiotics, antiviral agents, analgesics,
CC anesthetics, anti-inflammatories, antineoplastic agents and diagnostic
CC agents. LD optimize activity of PAD, e.g. by delivering it to selected
CC locations or by delivering it to neighboring cells (bystander effect),
CC and allow a reduction in dose of prodrug, and thus of systemic side-
CC effects. Nucleic acids encoding the agent may be expressed selectively
CC in hypoxic cells. The present sequence represents the single chain
CC variable antibody fragment against the tumor antigen 5T4 (5T4 scFv).
CC 5T4 scFv is used in the construction of a fusion protein comprising
CC 5T4 scFv and a human P450 reductase derivative alp450R.
XX
SQ Sequence 243 AA;

Query Match 27.6%; Score 972.5; DB 20; Length 243;
Best Local Similarity 40.1%; Pred. No. 6.9e-51;
Matches 227; Conservative 3; Mismatches 13; Indels 323; Gaps 5;

Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSFTGYMHVWVKQSPGKLEWIGRINPNNGVTLY 60
Db 1 EVQLQQSGPDLVKPGASVKISKASGYSFTGYMHVWVKQSHGKSLIEWIGRINPNNGVTLY 60
Qy 61 NQKFKDKATLVDKSSTTAYMELSLTSEDSAVYYCARSTMTITNYVMDYWGQGTSTVTVSS 120
Db 61 NQKFKDKAILTVDKSSTTAYMELSLTSEDSAVYYCARSTMTITNYVMDYWGQVTSVTYS- 119
Qy 121 AKTTPPSVYPLAPGSAQTNSNVTLGCLVKGYFPEPVTVTVNSGSLSSGVHTFPAVLQSD 180
Db 120 ----- 119
Qy 181 LYTLLSSVTVPSSTWPSSETVTCNVAHPASSTKVDDKIVPRDSGGPSEKSEINEKDLRKK 240
Db 120 -----SGG----- 122
Qy 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGCVTLHDNNRLTEKKVPINLWIDGKQ 360
Db 123 -----GGSGGGGT-----GG----- 132
Qy 361 TTVPIDKVKTSKKEVTVQELDLQARHYLHGKFLYNSDFGKVGORGLIVFHSSEGSTVS 420
Db 133 -----GG----- 134
Qy 421 YDLFDAQGGYPTDLLRIYRNDTTISSTLSLSLYLTTSIVMTQTPTSLYSAGDRVTIT 480
Db 135 -----SSIVMTQTPTSLYSAGDRVTIT 157
Qy 481 CKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVPDRFSGSGYGTDFLTITISVQAED 540
Db 158 CKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVPDRFISGSGYGTDFTTISTLQAEED 217
Qy 541 AAVYFCQQDYNSPPTFGGGTKLEIKR 566
Db 218 LAVYFCQQDYNSPPTFGGGTKLEIKR 243

RESULT 31
AAW86002
ID AAW86002 standard; Protein: 243 AA.
XX
AC AAW86002;
XX
DT 15-MAR-1999 (first entry)
XX
DE Murine anti-5T4 antigen monoclonal antibody scFv.
XX
KW Tumour interacting protein: cancer; gene therapy: vector;
KW 5T4 antigen; monoclonal antibody; single chain antibody; scFv;
XX mouse; 5T4scFv.1.
```

```
OS Chimeric - Mus sp.
OS Chimeric - synthetic.
FH Key Location/Qualifiers
FT Misc-difference 169 /note= "encoded by GDT"
XX
PN W09855607-A2.
XX
PD 10-DEC-1998.
XX
PF 04-JUN-1998; 98WO-GB01627.
XX
PR 04-JUL-1997; 97GB-0014230.
PR 04-JUN-1997; 97GB-0011579.
PR 20-JUN-1997; 97GB-0013150.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Bebbington CR, Carroll MW, Ellard FM, Kingsman SM;
PI Myers KA;
XX
DR WPI; 1999-059910/05.
DR N-PSDB; AAW80290.
XX
XX New vector encoding a tumour interacting protein for treating cancer
PT - contains a desired nucleotide sequence and/or protein which
PT recognises tumours, and is used as a gene delivery system to treat
PT cancer
XX
PS Example 1; Fig 1A; 82pp; English.
XX
CC This amino acid sequence comprises an scFv, termed 5T4scFv.1,
CC comprising the heavy chain variable region (VH) from the murine 5T4
CC monoclonal antibody followed by a 15-amino acid flexible linker and
CC the light chain variable region (VL) of the mouse 5T4 antibody. The
CC trophoblast cell surface antigen defined by monoclonal antibody 5T4
CC is expressed at high levels on the cells of a wide variety of human
CC tumours. 5T4scFv.1 DNA (see AAW80290) can be used to construct
CC single-chain antibodies (see AAW86003) and scFv fusion constructs
CC (see AAW86004-05). The invention relates to a vector comprising a
CC nucleotide sequence coding for a tumour interacting protein (TIP)
CC and optionally a nucleotide sequence of interest (NOI) which
CC encodes a protein of interest (POI), the vector being capable of
CC delivering the NOI and/or POI to the tumour recognised by the TIP.
CC Delivery can be in vivo or ex vivo. The vector is used to treat
CC cancer, and may also used as a gene delivery system for introducing
CC at least 1 gene encoding a TIP (preferably a tumour binding protein)
CC into a haematopoietic cell lineage.
XX
SQ Sequence 243 AA;
```

```
Query Match 27.6%; Score 972.5; DB 20; Length 243;
Best Local Similarity 40.1%; Pred. No. 6.9e-51;
Matches 227; Conservative 3; Mismatches 13; Indels 323; Gaps 5;

Qy 1 EVQLQQSGPDLVKPGASVKISKASGYSFTGYMHVWVKQSPGKLEWIGRINPNNGVTLY 60
Db 1 EVQLQQSGPDLVKPGASVKISKASGYSFTGYMHVWVKQSHGKSLIEWIGRINPNNGVTLY 60
Qy 61 NQKFKDKATLVDKSSTTAYMELSLTSEDSAVYYCARSTMTITNYVMDYWGQGTSTVTVSS 120
Db 61 NQKFKDKAILTVDKSSTTAYMELSLTSEDSAVYYCARSTMTITNYVMDYWGQVTSVTYS- 119
Qy 121 AKTTPPSVYPLAPGSAQTNSNVTLGCLVKGYFPEPVTVTVNSGSLSSGVHTFPAVLQSD 180
Db 120 ----- 119
Qy 181 LYTLLSSVTVPSSTWPSSETVTCNVAHPASSTKVDDKIVPRDSGGPSEKSEINEKDLRKK 240
Db 120 -----SGG----- 122
Qy 241 SELOQTALGNLKIYYNYSKAITSSSEKSAQDQFLNTLLFKGFFTGHWPWYNDLLVDLGSTA 300
```


KW light chain; heavy chain; Fab; monoclonal antibody; hypervariable region;
 KW infection.
 XX
 OS Mus sp.
 OS FR2777285-A1.
 PN
 PD 15-OCT-1999.
 XX
 XX 10-APR-1998; 98FR-0004876.
 XX
 XX 10-APR-1998; 98FR-0004876.
 XX
 XX (INMR) BIO MERIEUX.
 XX
 XX Novelli RA, Monaco S, Piga N, Berthet C, Mallet F, Cusack S;
 PI Chassaigne V;
 XX
 XX WPI: 1999-593428/51.
 DR N-PSDB; AA228805.
 XX
 XX New peptide ligand specific for p24 of human immune deficiency virus
 PT contains hypervariable regions of antibody 13B5, used for diagnosing
 PT HIV infection
 XX
 PS Claim 2: Page 20-21; 27pp; French.
 XX
 CC The invention relates to a peptide ligand with specific affinity for
 CC the p24 protein of human immune deficiency virus-1 (HIV-1) comprising
 CC at least one peptide strand corresponding to the N-terminal region of
 CC the light and/or heavy chain of the Fab fragment of monoclonal antibody
 CC 13B5 in which: (i) the light chain includes three hypervariable regions
 CC (HVR) at amino acid (aa) positions 24-33, 49-55 and 88-95 of AAY44175;
 CC and (ii) the heavy chain includes three HVR at aa positions 26-35,
 CC 49-65 and 99-109 of this sequence. The peptide ligands are reagents
 CC for detecting p24 (by standard immunoassays) in biological samples,
 CC specifically for diagnosis of HIV-1 infection or can be used to treat
 CC HIV-1 infections.
 XX
 SQ Sequence 214 AA:
 Query Match 27.5%; Score 970; DB 20; Length 214;
 Best Local Similarity 85.0%; Pred. No. 8.3e-51;
 Matches 182; Conservative 15; Mismatches 17; Indels 0; Gaps 0;
 QY 1 EVLOQSGDPLVKPGASVKISCKASYSTGYMHVWVKSPGKLEWIGRIINPNNGVTLY 60
 DB 1 EVLOQSGAELARPGASVKMSCKASYTFTSYMHVVKRPGOGLEWIGYINPSSGYSNY 60
 QY 61 NQKFKDKATLTVDKSSSTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGQGTSTVTVSS 120
 DB 61 NQKFKDKATLTADKSSSTAYMQLSSLTSDSAVYYCSRVPVRLGYNFYWGQGTSTLTSS 120
 QY 121 AKTTPPSVYPLAPGSAQAQNSVMTLGCLVKGYFPEPTVTWNSSGSLSSGVHTFPAVLQSD 180
 DB 121 AKTTPSVYPLAPGSAQAQNSVMTLGCLVKGYFPEPTVTWNSSGSLSSGVHTFPAVLQSD 180
 QY 181 LYTSSSVTPSPSTWSPSETVTCNVAHPASSTKVD 214
 DB 181 LYTSSSVTPSPSTWSPSEIVTCNVAHPASSTKVD 214
 RESULT 34
 ID AA086004
 XX AA086004 standard; Protein; 488 AA.
 AC AA086004;
 XX
 XX 15-MAR-1999 (first entry)
 XX Human B7-1.5T4.1 protein fusion, specific for human 5T4.
 XX Tumour interacting protein; cancer; gene therapy; vector;
 KW

KW 5T4 antigen; monoclonal antibody; single chain antibody;
 KW mouse; human; B7-1; co-stimulatory molecule.
 XX
 OS Chimeric - Mus sp.
 OS Chimeric - Homo sapiens.
 OS Chimeric - synthetic..
 XX
 PN WO9855607-A2.
 XX
 XX 10-DEC-1998.
 XX
 XX 04-JUN-1998; 98WO-GB01627.
 PF
 XX 04-JUL-1997; 97GB-0014230.
 PR 04-JUN-1997; 97GB-0011579.
 PR 20-JUN-1997; 97GB-0013150.
 XX
 XX (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 XX Bebbington CR, Carroll MW, Ellard FM, Kingsman SM;
 PI Myers KA;
 XX
 XX WPI: 1999-059910/05.
 DR N-PSDB; AAV80292.
 XX
 XX New vector encoding a tumour interacting protein for treating cancer
 PT - contains a desired nucleotide sequence and/or protein which
 PT recognises tumours, and is used as a gene delivery system to treat
 PT cancer
 XX
 XX Example 5; Fig 2; 82pp; English.
 PS
 CC This is the amino acid sequence of B7-1.5T4.1, a fusion protein
 CC comprising the extracellular domain (amino acids 1-215) of human
 CC co-stimulatory molecule B7-1 joined via a flexible peptide linker
 CC to an scFv (see AAW86002) derived from murine 5T4 monoclonal
 CC antibody. B7-1.5T4.1 cDNA (see AAV80292) can be inserted into vector
 CC pCI to allow expression of the fusion protein in mammalian cells.
 CC The trophoblast cell surface antigen defined by 5T4 is expressed at
 CC high levels on the cells of a wide variety of human tumours. The
 CC invention relates to a vector comprising a nucleotide sequence
 CC coding for a tumour interacting protein (TIP) and optionally a
 CC nucleotide sequence of interest (NOI) which encodes a protein of
 CC interest (POI), the vector being capable of delivering the NOI
 CC and/or POI to the tumour recognised by the TIP. Delivery can be in
 CC vivo or ex vivo. The vector is used to treat cancer, and may also
 CC used as a gene delivery system for introducing at least 1 gene
 CC encoding a TIP (preferably a tumour binding protein) into a
 CC haematopoietic cell lineage. B7-1 is expected to bind specifically
 CC to CD28 and CTLA-4 present on human T-cells.
 XX
 SQ Sequence 488 AA:
 Query Match 27.5%; Score 967.5; DB 20; Length 488;
 Best Local Similarity 40.0%; Pred. No. 3.2e-50;
 Matches 226; Conservative 3; Mismatches 13; Indels 323; Gaps 5;
 QY 1 EVLOQSGDPLVKPGASVKISCKASYSTGYMHVWVKSPGKLEWIGRIINPNNGVTLY 60
 DB 247 EVLOQSGDPLVKPGASVKISCKASYSTGYMHVWVKSHGKSLWIGRIINPNNGVTLY 306
 QY 61 NQKFKDKATLTVDKSSSTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGQGTSTVTVSS 120
 DB 307 NQKFKDKAILTVDKSSSTAYMELRSLTSDSAVYYCARSTMITNYYMDYWGQVTSTVTS- 365
 QY 121 AKTTPPSVYPLAPGSAQAQNSVMTLGCLVKGYFPEPTVTWNSSGSLSSGVHTFPAVLQSD 180
 DB 366 -----SGG----- 365
 QY 181 LYTSSSVTPSPSTWSPSETVTCNVAHPASSTKVDKIVPRDSGGPSEKSEINEKDLRKK 240
 DB 366 -----SGG----- 368

QY 241 SELQGTALGNLQIYYNSKAITSEKSAQDLTNTLLFKGFFTHPWNVDLLVLDGSTA 300
 Db ----- 368
 QY 301 ATSEYEGSSVDLYGAYYQACAGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
 Db ----- 378
 QY 361 TTVPIDKVKTSKKEVTQVQLDQARHYLHGKFGLYNSDSFGKVGQGLVHFHSSGEGSTVS 420
 Db ----- 380
 QY 421 YDLFDAQGOYPTDILLRIYRDNTTISSTLSLSLYLYTTISIVMTQPTSLLSVAGDRVTIT 480
 Db ----- 403
 QY 481 CKASQSVNDVAMVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTISSVQAE 540
 Db ----- 463
 QY 541 AAVYFCQDDYNSPPTFGGCTKLEIK 565
 Db ----- 488
 QY 464 LAVYFCQDDYNSPPTFGGCTKLEIK 488
 Db -----
 RESULT 35
 AAB83836
 ID AAB83836 standard; Protein; 488 AA.
 AC AAB83836;
 XX
 DT 23-JUL-2001 (first entry)
 DE Amino acid sequence of a B7-1.5T4.1 fusion protein.
 XX
 KW Single chain antibody; ScFv; inflammatory disease; arthritis; cancer;
 KW hypersensitivity; autoimmune disease; central nervous system disorder;
 KW Parkinson's disease; periodontal disease; cardiopulmonary disease;
 KW cardiovascular disease; gastrointestinal disease; infection; diabetes;
 KW Helicobacter-related disease; immune disorder.
 XX
 OS Synthetic.
 OS Mus sp.
 OS Homo sapiens.
 XX
 PN WO200136486-A2.
 XX
 PD 25-MAY-2001.
 XX
 PF 13-NOV-2000; 2000WO-CB04317.
 XX
 PR 18-NOV-1999; 99WO-CB03859.
 PR 15-FEB-2000; 2000GB-0003527.
 PR 02-MAR-2000; 2000GB-0005071.
 XX
 PA (OXFO-) OXFORD BIOMEDICA UK LTD.
 XX
 PI Kingsman A, Kingsman SM, Bebbington CR, Carroll MW, Ellard PM;
 PI Myers KA;
 XX
 DR WPI; 2001-343805/36.
 DR N-PSDB; AAF89730.
 XX
 PT Use of single chain antibody capable of recognizing a disease
 PT associated molecule for manufacturing a medicament for preventing
 PT and/or treating a disease condition associated with disease associated
 PT molecule -
 XX
 PS Claim 3; Fig 2; 118pp; English.
 PS
 CC The specification describes the use of a single chain antibody (ScFv),
 CC which is capable of recognizing a disease associated molecule in the
 CC manufacture of a medicament for the prevention and treatment of a

CC disease condition. The ScFv antibody is useful in the manufacture of
 CC a medicament, for affecting a disease in vivo, for preparing a
 CC pharmaceutical composition, for in vivo imaging and/or for adjuvant
 CC treatment of a disease. The ScFv antibody is also useful for
 CC treating inflammatory diseases including arthritis, hypersensitivity,
 CC autoimmune diseases, cancers, central nervous system disorders
 CC including Parkinson's disease, periodontal diseases, cardiopulmonary
 CC diseases, cardiovascular diseases, gastrointestinal disorders,
 CC infections, diabetes, Helicobacter-related diseases, and other, immune
 CC disorders. The present sequence represents a B7-1.5T4.1 fusion protein.
 CC This comprises the N-terminus of the 5T4 ScFv is fused after amino acid
 CC 215 of human B7-1.
 XX
 SQ Sequence 488 AA;
 Query Match 27.5%; Score 967.5; DB 22; Length 488;
 Best Local Similarity 40.0%; Pred. No. 3.2e-50;
 Matches 226; Conservative 3; Mismatches 13; Indels 323; Gaps 5;
 QY 1 EVLOQSGDLYKPGASVKISCKASGYSTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60
 Db ||||| 323
 Db 247 EVLOQSGDLYKPGASVKISCKASGYSTGYIMHWKQSHGKSLWIGRINPNNGVTLY 306
 QY 61 NOKFKDKATFTVDKSSTTAYMELSLTSDSAVYYCARSTMITNYMDYWGQCTSVTVSS 120
 Db ||||| 365
 Db 307 NOKFKDKATFTVDKSSTTAYMELSLTSDSAVYYCARSTMITNYMDYWGQCTSVTVSS- 365
 QY 121 AKTTPPSVYPLAPGAAQNTNSWTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
 Db ----- 365
 QY 181 LYTSSSVTPSPSTWPTETVCNVAHPASSTKVKKIVPRDSGGPSEKSEEINEXDLRKK 240
 Db ----- 368
 QY 241 SELQGTALGNLQIYYNSKAITSEKSAQDLTNTLLFKGFFTHPWNVDLLVLDGSTA 300
 Db ----- 368
 QY 301 ATSEYEGSSVDLYGAYYQACAGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDGKQ 360
 Db ----- 378
 QY 361 TTVPIDKVKTSKKEVTQVQLDQARHYLHGKFGLYNSDSFGKVGQGLVHFHSSGEGSTVS 420
 Db ----- 380
 QY 421 YDLFDAQGOYPTDILLRIYRDNTTISSTLSLSLYLYTTISIVMTQPTSLLSVAGDRVTIT 480
 Db ----- 403
 QY 481 CKASQSVNDVAMVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTISSVQAE 540
 Db ----- 463
 QY 541 AAVYFCQDDYNSPPTFGGCTKLEIK 565
 Db ----- 488
 Db 464 LAVYFCQDDYNSPPTFGGCTKLEIK 488
 Db -----
 RESULT 36
 AARL3061
 ID AARL3061 standard; Protein; 468 AA.
 XX
 AC AARL3061;
 XX
 DT 03-OCT-1991 (first entry)
 XX
 DE Monoclonal antibody OK3T heavy chain.
 XX
 KW OK3T; light chain; humanised antibodies; CDR-grafting.
 XX
 OS Mus musculus.

CC	forms of the antigen. The antibodies generated can be used in the
CC	diagnosis, e.g. detection of the immunorecessive antigen, or in therapy
CC	e.g. of cancer, Alzheimer's disease or familial hypercholesterolaemia.
CC	The method of production of the antibody allows rapid and sensitive
CC	isolation of antibodies that would be difficult to isolate by standard
CC	methods. The antibodies produced have greater binding affinity than
CC	those produced by combinatorial/hybridoma methods.
XX	
SQ	Sequence 223 AA;
	Query Match 27.1%; Score 953.5; DB 16; Length 223;
	Best Local Similarity 83.08; Pred No. 8.6e-50;
	Matches 185; Conservative 12; Mismatches 23; Indels 3; Gaps 2;
Qy	1 EVQLQQSGPDILVKPGASVKISKCKASGYSTFTGYYMHVWKQSPKGLEWIGRINPNNGVTLY 60
Dd	1 EVQLQQSGPELMPFGASVKISKCKATGLTSSYLEWYKQSPGHGLEWIGEITLFCGSAHY 60
Qy	61 NQKFKDKATLVDXKSSTAYMELSLTSSEDSAVYYCARSTMITNY--VMDYWGOGTSTVTV 118
Dd	61 NEKPKGKATFVDTSNNTAYMQLSLTSEDSDAVYYCARGD-YGNYGDFDYWGOGTTLTIV 119
Qy	119 SSAKTTPPSVPLAPGSAAOQNSMTVLGCLVKGYFEPETVTTWNSGSLSSGVHTFPVLIQ 178
Dd	120 SSAKTTPPSVPLAPGSAAOQNSMTVLGCLVKGYFEPETVTTWNSGSLSSGVHTFPVLIQ 179
Qy	179 SDLYTLSSSVTPSSTWPSETVTCNVNHPASTKVDDKIVPRD 221
Dd	180 SDLYTLSSSVTPSSTWPSETVTCNVNHPASTKVDDKIVPRD 222
AC	AAR53802;
DT	23-DEC-1994 (first entry)
DE	FAB light chain for IMCP.
KW	McPC603; V-min; synthetic gene; D1.3; antibody; immunotoxins; tumour;
KW	intracellular expression; E. coli; protease deficient; treatment;
KW	ligand-binding variable domain; framework region; beta-barrel;
KW	beta-sheet; linker; complementarity determining regions; radioimaging;
KW	prodrug therapy; diagnosis.
OS	Synthetic.
XX	
PN	WO9412625-A.
XX	
PD	09-JUN-1994.
XX	
PF	19-NOV-1993; 93WO-GB02375.
XX	
PR	23-NOV-1992; 92GB-0024588.
PR	31-DEC-1992; 92GB-0027189.
XX	
PA	(ZENE) ZENECA LTD.
PI	
PI	Slater AM, Timms D;
DR	
DR	WPI; 1994-200254/24..
XX	
PT	New ligand variable binding domains of reduced size - contain
PT	complementarity determining regions and framework region to
PT	provide correct orientation, also related DNA, expression vectors
PT	etc., useful for radio-imaging and tumour treatment
XX	
FS	Disclosure; Fig 4; 187pp; English.
XX	
CC	The sequences given in AAR53802-05 represent the FAB light and heavy
CC	chains of F19.9 and IMCP. These FAB had the V-min protein ligated

CC into them to form a conjugate. V-min is a ligand-binding variable
CC domain which comprises a framework region of a cyclically permuted
CC central beta-barrel, outer beta-sheet segments, and linker segments,
CC and complementarity determining regions. V-min, when conjugated, are
CC useful in radioimaging, as immunotoxins and in antibody directed
CC enzyme prodrug therapy, ie. for diagnosis or treatment of tumours.
XX
SQ Sequence 220 AA:

Query Match 27.1%; Score 953; DB 15; Length 220;
Best Local Similarity 84.4%; Pred. No. 9e-50;
Matches 184; Conservative 14; Mismatches 14; Indels 6; Gaps 1;
QY 460 IYVOTPTSLVLSAGDRTVITTCASQVSNVND-----VAWYQKPGQSPKLLISYTSRRY 513
DB 2 IYVOTGSPSLVLSAGERTVMSKSSQSLNSQNFNLAWYQKPGQSPKLLIYGASTRE 61
QY 514 AGVPRDFSGSGYGTDTLTISVQAEADAAYFCQDYNPPTFGGKLEIKRADAAPT 573
DB 62 SGVPRDFSGSGYGTDTLTISVQAEADLAAYVCQNDHSYPLTFGAGTKLEIKRADAAPT 121
QY 574 SIFFPSSEQLTSGGASVVCFLNFPKIDNVKWKIDGSRQNGVLSNWTQDQSKDSTYSM 533
DB 122 SIFFPSSEQLTSGGASVVCFLNFPKIDNVKWKIDGSRQNGVLSNWTQDQSKDSTYSM 181
QY 634 SSTLTITKDEYERHNSYTCETHKTSPIVKSFNRRNE 671
DB 182 SSTLTITKDEYERHNSYTCETHKTSPIVKSFNRRNE 219

RESULT 42
AAR40384
ID AAR40384 standard; Protein; 469 AA.
XX
AC AAR40384;
XX
DT 08-FEB-1994 (first entry)
DE Monoclonal antibody M(alpha)2-3 Heavy-chain.
XX
XX anti-snake small neurotoxin antibody; heavy chain; IgG2;
KW immunoglobulin; bispecific bivalent antibody; cell-targetting;
KW cytotoxic agent.

XX Key Location/Qualifiers
FT Peptide 1..19
FT Region /label= signal_peptide
FT Region 20..139
FT Region /label= variable
FT Region 140..236
FT Region /label= constant
FT Region 237..252
FT Region /label= joining
FT Region 253..362
FT Region /label= constant
FT Region 363..469
FT Region /label= constant
XX
PN EP556111-A.
XX
PD 18-AUG-1993.
XX
PF 09-FEB-1993; 93EP-0400323.
XX
PR 11-FEB-1992; 92FR-0001505.
XX
PA (BOUL/) BOULAIN J.
XX (COMS) COMMISSARIAT ENERGIE ATOMIQUE.
XX
PI Boulain J, Ducancel F, Gillet D, Menez A;
XX
XX WPI: 1993-260351/33.
DR N-PSDB; AAQ48037.

XX New immunoglobulin hybrid proteins - with immunoglobulin
PT fragments linked to dimeric protein, for diagnostic or
PT therapeutic use
XX
XX Example 1; Fig 3A; 37pp; French.
XX
CC A fragment of the heavy chain (VH + CH1) from the anti-snake small
CC neurotoxin monoclonal antibody M(alpha)2-3 was PCR-amplified from
CC hybridoma-derived cDNA using primers AAQ48039 and AAQ48040. A light
CC chain fragment (VL + CL) was amplified from the same source using
CC primers AAQ48041 and AAQ48042. The two amplified fragments were
CC inserted into the same vector; the H-chain fragment was inserted
CC (in-frame) between codons 6-7 of the phoA coding sequence and the
CC L-chain fragment was inserted into a cassette which contained a
CC phoA 5'-3' sequence, a signal peptide and the first 6 codons of phoA.
CC The cassette was positioned between the termination codon and
CC the transcription termination sequence of phoA. The fusion
CC construct is expected to encode a hybrid protein comprising two
CC identical Ab-derived units. The invention also covers hybrid
CC proteins containing two different Ab-derived units (i.e. to produce
CC bispecific antibodies). When a toxic protein is used in place of
CC phoA, the hybrid molecules can be used as cell-targetting
CC therapeutic agents.

XX Sequence 469 AA:
SQ
Query Match 27.0%; Score 951.5; DB 14; Length 469;
Best Local Similarity 34.2%; Pred. No. 2.8e-49;
Matches 236; Conservative 62; Mismatches 126; Indels 267; Gaps 13;
QY 1 EVLOQSGPDLVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGRINPNNGVTLY 60
DB 20 QIQLOQSGPELVKPGASVKISKASGYSTGYIMHWKQSPGKLEWIGRINPNNGVTLY 79
QY 61 NQKFKDKATLVKDSSTAYMELRLSLTSEDSAVYYCARSTMTINTYMDYWGOGTSVTYSS 120
DB 80 NENFKGKATLVDTSSSTAYMQLSLSLTSSEDTAVYFCARAMGATATLLDYWGOGTTLTVSS 139
QY 121 AKTTPSYVPLAPGSAQAOTNSMTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
DB 140 AKTTAPSYVPLAPVCGDGTGSSVTGLCLVKGYFPEPVTLTWNSGSLSSGVHTFPAVLQSD 199
QY 181 LYTSSSVTPSPSTWPSSETVTCNVAHPASSTKVDKKIIPRDSGGPSEKSEINEKDLRKK 240
DB 200 LYTSSSVTPSTWPSSETVTCNVAHPASSTKVDKKIIPRDSGGPSEKSEINEKDLRKK 247
QY 241 SELOCTALGNLKOIYYNYSKAITSEKSAQDLTNTLLFKGFTGHPWYNDLLDLGSTA 300
DB 248 ----- 247
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTCAGYGVTLHDNNRLTEKKVPINLWIDGKQ 360
DB 248 -----PCKCPAPN-----LLGSPSVF----- 263
QY 361 TTVPIDKVKTSKKEVTQVQLDQARHLYLHGKFLYNSDSFGKGVQGLIVFHSSEGSVTS 420
DB 264 -----IFPPKIKDVLMI---SLSPITV 282
QY 421 YDLFDAQGYPTLLRIYRDNTTISSTLSISLYLYTTISIVMTQPTTLLVSAGDRVTIT 480
DB 283 CVVVDVSEDDPD-----VQISWFFVNNVEVHTAQOT----- 313
QY 481 CKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVPDRFSGSGYGTDTLTISVQAEQ 540
DB 314 -----HREDYNSLRVY----- 325
QY 541 AAVYFCQDYNPPTFGGKLEIKRAD-----AAPTVISIFPPSSEQLT 584
DB 326 SALPTQHQDWSGKEF-----KCKVNNKDLPAPIERTISKPGSVRAPQVYVLPPEEEMT 381
QY 585 SGGASVVCFLNFPKIDNVKWKIDGSRQNGVLSNWTQDQSKDSTYSMTLTL 639
DR N-PSDB; AAQ48037.

Db	382	KKQVTLTCMTDFMPEDIIYVETWNGNGTELNYKNTPEVLDS-----DGSYPMYKSLRV	430
Qy	640	TKDEYERHNSYTCSEATHK-TSTSPIVKSFNR	669
Db	435	EKKWVERNYSYSCSVVHEGLNHHHTTKSFSR	465
RESULT 43			
AAU78253	ID	AAU78253 standard; Protein; 214 AA.	
XX	AC	AAU78253;	
XX	DT	05-JUN-2002 (first entry)	
XX	XX	Mouse agglutination antibody FAB region gene associated protein #1.	
DE	DE	FAB region; gene; mouse; agglutination monoclonal; antibody; agonist;	
KW	KW	human blood cell.	
KW	XX	Unidentified.	
OS	XX	KR99048543-A.	
PN	PN	05-JUL-1999.	
PD	PD	10-DEC-1997; 97KR-0067281.	
PF	PF	10-DEC-1997; 97KR-0067281.	
XX	XX	(KOAD) KOREA ADV INST SCI & TECHNOLOGY.	
PA	PA	Lim SY, Cho SH, Lee YI, Bong YS;	
PI	PI	WPI; 2000-449017/39.	
XX	XX	FAB region gene of mouse agglutination monoclonal antibody against	
PT	PT	human blood cell and base sequence thereof -	
PT	PT	Disclosure; Page 6; 9pp; Korean.	
PS	PS	The present invention relates to a new Fab region gene of mouse	
XX	XX	agglutination monoclonal antibody against human blood cell. The prese	
CC	CC	amino acid sequence represents the associated protein sequence #1 tha	
CC	CC	was used in the methods of the invention.	
XX	XX	Sequence 214 AA;	
SQ	SQ	Query Match 26.9%; Score 948; DB 21; Length 214;	
		Best Local Similarity 85.8%; Pred. No. 1.7e-49;	
		Matches 182; Conservative 9; Mismatches 21; Indels 0; Gaps	
Qy	460	IVMTQTPTSLVLSAGDRVTTTCRKASQSVNDVAMVQKPGQSPKLLISYTSRYAGVPDR	519
Db	2	IVMTQSPKFMSTVGDRAVSVTCASQIVGTGNVAMVQKPGQSPKALIYSASYRYSGV	61
Qy	520	FSGSGYCTDFTLTLSISVQAEADAAVFCQQDYNPPTFGGCTKLEIKRADAAPT	579
Db	62	FTGSGTGTDTLTLSINVOSEDAEYFCQQYRNPYPTFGGCTKLEIKRADAAPT	121
Qy	580	SEQLTSGGASVVCFLNNFYPKDINVKWKIDGSRQNGVLNSWTDQDSKDYSMSTLTL	639
Db	122	SEQLTSGGASVVCFLNNFYPKDINVKWKIDGSRQNGVLNSWTDQDSKDYSMSTLTL	181
Qy	640	TKDEYERHNSYTCSEATHKTSPIVKSFNRE	671
Db	182	TKDEYERHNSYTCSEATHKTSPIVKSFNRE	213
RESULT 44			
AAU14104	ID	AAU14104 standard; peptide; 257 AA.	
XX	XX		

[illegible]

	QY	400	RGLVVFHSSEGSIVSIDLFDAGQQGYPDTLLRLIKRNTIISSISLSLSLEYIT	2578
	QY	400	IIIIIII:I :IIIIII IIII I IIIIII IIII I :IIIIII	2579
Dbb		205	RGLIVFHTSTEPSNYDLEGAQQGOYSNTLLRIYDRNKTINSENMMIDIIYLKTS	2580

RESULT 45	
AA76087	
ID	AA76087 standard; Protein; 239 AA.
XX	
XX	AA76087;
XX	
DT	21-NOV-1995 (first entry)
XX	
DE	MAB 55.1 light chain.
XX	
KW	Antigen binding structure; complementarity
KW	CA55.1; colorectal cancer; tumor-associated
KW	monoclonal antibody; MAb; immunotherapy;
KW	transgenic animal; transgenic plant; anti-
KW	humanized antibody; immunotoxin.
XX	
OS	Mus sp.
XX	
FH	Key
FT	Peptide
FT	1..20
FT	/label= Sig_peptide
FT	21..239
FT	/label= Mat_protein
FT	/note= "claim 3, page 98"
XX	
PN	WO9515382-A.
XX	
PD	08-JUN-1995.
XX	
PF	29-NOV-1994; 94WO-GB02610.
XX	
PR	03-JUN-1994; 94GB-0011089.
PR	03-DEC-1993; 93GB-0024819.
XX	
XX	(ZENE) ZENECA.LTD.
XX	
PI	Blakey DC, Boot C, Copley CG, Hall SM,
PI	Rose MS, Wright AF;
XX	
DR	WPI; 1995-215262/28.
DR	N-PSDB; AAQ94036.
XX	
PT	Antigen binding structures containing CD8
PT	antigen - produced by hybridomas and host
PT	diagnosis and therapy of cancer
XX	
PS	Disclosure; Fig.16; 12lpp; English.
XX	
CC	MAB 55.1 (ECACC 93081901) recognises the
CC	antigen CA55.1. cDNAs for the heavy (AAQ
CC	chains of 55.1 were isolated, and F(ab') ₂
CC	v-mim humanized 55.1 constructs have been
CC	cell and E. coli.
XX	
SQ	Sequence 239 AA;

D6	182	NGVINSWTDDSKDSTYSMSLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRNE	238
<p> </p>			
RESULT 46			
AAW06738	ID	AAW06738 standard; Protein: 233 AA.	
XX	AC	AAW06738;	
XX	DT	08-MAR-1997 (first entry)	
XX	XX	Staphylococcus enterotoxin A.	
DE	XX	Enterotoxin A; superantigen; antigen; cytokine; chemokine; T cell;	
KW	KW	lymphocyte; monocyte; natural killer cell; gene therapy; cancer;	
KW	KW	vaccine; adjuvant.	
XX	OS	Staphylococcus sp.	
XX	PN	W09636366-A1.	
XX	PD	21-NOV-1996.	
XX	XX	20-MAY-1996; 96WO-US07432.	
Pf	PF	29-DEC-1995; 95US-0508006.	
PR	PR	18-MAY-1995; 95US-0446918.	
XX	XX	(NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY MED.	
PA	PI	Dow SW, Elmslie RE, Potter TA;	
XX	XX	WPI: 1997-011857/01.	
DR	DR	N-PSDB: AAT45699.	
XX	PT	Recombinant molecule encoding superantigen and opt. cytokine or	
PT	PT	chemokine - controls activity of effector cells (T cells, monocytes,	
PT	PT	natural killer cells), used for gene therapy of cancer	
XX	PS	Example 1; Page 98-99; 131pp; English.	
XX	CC	A cDNA clone (AAT45699) codes for staphylococcal enterotoxin A	
XX	CC	(AAW06738) superantigen. Nucleic acids encoding superantigens (see	
CC	CC	also AAW06737, AAW06739), esp. truncated forms of the superantigen	
CC	CC	lacking the leader peptide, can be used in the gene therapy of	
CC	CC	cancer, infectious diseases and immunological disorders. The	
CC	CC	nucleic acid, optionally in combination with cytokine or chemokine	
CC	CC	nucleic acids, is delivered to an animal using e.g. liposomes. It	
CC	CC	acts by controlling the activity of effector cells, such as T-cells,	
CC	CC	macrophages, monocytes and/or natural killer cells. Localised	
CC	CC	prodn. of an effective but non-toxic amount of encoded proteins	
CC	CC	allows safe treatment of the animal.	
XX	XX	Sequence 233 AA:	
S0			

Db 182 GLIVFHTSTEPSVNYDLFGAOGQVSNTLLRIYRDNKTINSENMHIDIYLYTS 233

RESULT 47
AAR45011
ID AAR45011 standard; protein; 233 AA.
XX
XX AAR45011;
XX
XX 08-JUN-1994 (first entry)
XX
XX Staphylococcal enterotoxin SEA.
XX
XX Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
XX autoimmune disease; toxicity; Protein A; perfusion system.
XX
XX Staphylococcus aureus.
XX
XX
XX Key Location/Qualifiers
XX Misc-difference 49 /note= "Given in the specification as O, no further
XX details given"
XX
XX WO9324136-A.
XX
XX 09-DEC-1993.
XX
XX 01-JUN-1993; 93WO-US05213.
XX
XX 01-JUN-1992; 92US-0891718.
XX
XX (STON/) STONE J L.
XX (TERM/) TERMAN D S.
XX
XX Stone JL, Terman DS;
XX
XX WPI; 1993-405418/50.
XX
XX Use of staphylococcal enterotoxin(s) and homologues - for
XX treating cancer in a patient or for the treatment of auto-immune
XX diseases
XX
XX Disclosure; Fig 1; 90pp; English.
XX
XX The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
XX which may be used in the methods of the invention for treating cancer
XX in a patient. These SEs, and homologues of them, can be used as
XX tumouricidal agents for treating cancers and autoimmune disease.
XX They exhibit tumouricidal activity and toxicity identical to that
XX observed for the Protein A perfusion system. They may be administered
XX by i.v. injection.
XX
XX
XX Sequence 233 AA;
Query Match 26.7%; Score 942; DB 14; Length 233;
Best Local Similarity 76.0%; Pred. No. 4.4e-49;
Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;
Qy 226 SEKSEINEKDLRKSELOGTALGNLKIYYNKAITSSEKSAQFLTNTLLFKGFTG 285
Db 1 SEKSEINEKDLRKSELOGTALGNLKIYYNKAITSSEKSAQFLTNTLLFKGFTG 60
Qy 286 HPWNLLDLVLSGTAANTSEYESSVDLYGAYGYOCAGTGNKTACMYGGVTLHDNNBLT 345
Db 61 HSWNDLLDFDSKDIDYKGGKVDLYGAYGYOCAGTGNKTACMYGGVTLHDNNBLT 120
Qy 346 BEKKVPINLWDGKQTTVPIDKVKTSKREVTAVQELDLQARHVLHGKGLYNSDSFGGKVQ 405
Db 121 BEKKVPINLWDGKQTTVPLETVTNKNVTVQELDPAARYLQEKYLNLYNSDVPDGKVQ 180
Qy 406 RGLIVFHTSTEPSVNYDLFGAOGQVPTLLRIYRDNKTINSENMHIDIYLYTT 458

Db 181 RGLIVFHTSTEPSVNYDLFGAOGQVSNTLLRIYRDNKTINSENMHIDIYLYTS 233

RESULT 48
AAR76086
ID AAR76086 standard; Peptide; 219 AA.
XX
XX AAR76086;
XX
XX 21-NOV-1995 (first entry)
XX
XX Mab 55.1 light chain.
XX
XX
XX Antigen binding structure; complementarity determining region; CDR;
XX CA55.1; colorectal cancer; tumor-associated antigen; hybridoma;
XX monoclonal antibody; Mab; immunotherapy; therapy; diagnosis;
XX transgenic animal; transgenic plant; antibody engineering;
XX humanized antibody; immunotoxin.
XX
XX Mus sp.
XX
XX WO9515382-A.
XX
XX 08-JUN-1995.
XX
XX 29-NOV-1994; 94WO-GB02610.
XX
XX 03-JUN-1994; 94GB-0011089.
XX 03-DEC-1993; 93GB-0024819.
XX
XX (ZENE) ZENECA LTD.
XX
XX Blakey DC, Boot C, Copley CG, Hall SM, Paterson DS;
XX Rose MS, Wright AF;
XX
XX WPI; 1995-215262/28.
XX
XX Antigen binding structures containing CDRs recognising the CA55.1
XX antigen - produced by hybridomas and host cells, for use in the
XX diagnosis and therapy of cancer
XX
XX Claim 3; Page 98; 121pp; English.
XX
XX An antigen binding structure is based on the CDRs (given in AAR76078-
XX 84) of the heavy (AAR76085) and light (AAR76086) chains of Mab 55.1
XX (EACC 93081901), which recognises the colorectal tumor-associated
XX antigen CA55.1. It is optionally humanized and in the form F(ab')₂,
XX F(ab)', Fab, Fv, scFv or V-min, and is produced in transgenic
XX animals or plants.
XX
XX Sequence 219 AA;
Query Match 26.7%; Score 941.5; DB 16; Length 219;
Best Local Similarity 83.0%; Pred. No. 4.4e-49;
Matches 181; Conservative 19; Mismatches 11; Indels 7; Gaps 2;
Qy 460 IVMTOTPTSLVSGADRYTITCKASQVSND-----VAVYQOKPGQSPKLLISYTSRY 513
Db 2 IVMSQPSLAVSAGEKVTNCKSSQSLNSTRKNYLAWTQQRPQSPKLLIYWASTRT 61
Qy 514 AGVPDRFSGSGYGTDTLTITSSVQAEDAAYFCQODYNSPPTFGGCTKLEIKRADAAPT 573
Db 62 SGVPDRFSGSGYGTDTLTITSSVQAEDLAIYCKQSY-TLRTFGGCTKLEIKRADAAPT 120
Qy 574 SIFPPSSQLTSGGASVVCFLNNFYPKDIINVKWKIDGSEKRONGLVNSWTDQSKDSTYSM 633
Db 121 SIFPPSSQLTSGGASVVCFLNNFYPKDIINVKWKIDGSEKRONGLVNSWTDQSKDSTYSM 180
Qy 634 SSTLTTLTKDEYERHNSYTCETHKTSTSPIVKSFNRNE 671
Db 181 SSTLTTLTKDEYERHNSYTCETHKTSTSPIVKSFNRNE 218

RESULT 49
AAW35373
ID AAW35373 standard; peptide; 233 AA.
XX
AC AAW35373;
XX
DT 20-APR-1998 (first entry)
XX
DE Staphylococcus enterotoxin SEA wild-type superantigen.
XX
KW SEE; SEA; Staphylococcus enterotoxin; superantigen; conjugate;
KW treatment; cancer; infection; autoimmune disease; antibody.
XX
OS Staphylococcus sp.
XX
PN W09735932-A1.
XX
PD 09-OCT-1997.
XX
PF 26-MAR-1997; 97WO-SE00537.
XX
PR 12-AUG-1996; 96US-0695692.
PR 29-MAR-1996; 96SE-0001245.
XX
PA (PHAA) PHARMACIA & UPJOHN AB.
XX
PI Abrahmsen L, Antonsson P, Bjoerk P, Dohlisten M;
PI Forsberg G, Hansson J, Kalland T;
XX
DR WPI; 1997-503052/46.

Conjugate of target seeking moiety and modified superantigen -
useful for activating the immune system to treat cancer, viral
infections, parasitic infestations and autoimmune diseases

Claim 8; Pages 36-37; 58pp; English.

This is the wild-type Staphylococcus enterotoxin SEA superantigen. This
SEA superantigen can be modified to be used in a novel conjugate. The
novel conjugate comprises a target seeking moiety and a modified wild
type superantigen. The modified superantigen retains its ability to
activate a subset of T cells, even though 1 or more wild-type amino acid
residues in at least 1 region which functions in determining binding to
T cell receptor (TCR) and activation of a subset of T cells has/have been
replaced. Such a modified superantigen can optionally be used as part of
a conjugate with a target seeking moiety, for activating the immune
system to treat a mammalian disease. A pharmaceutical composition can be
prepared comprising a modified antibody (preferably a Fab fragment fused
to a peptide moiety providing activation of T cells in Vbeta specific
manner) in which cysteines providing for interchain cysteine linkages in
the native antibody have been replaced (preferably by serine residues) to
prohibit cysteine formation. The modified wild-type superantigen is used
for treating cancer, viral infections, parasitic infestations and
autoimmune disease. The modified wild type superantigen has a lower
immunogenicity and reactivity with neutralising antibodies and has fewer
side-effects when used as a drug, compared to wild type superantigen.

Sequence 233 AA;

Query Match 26.7%; Score 941; DB 18; Length 233;
Best Local Similarity 76.0%; Pred. No. 5,1e-49;
Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

QY 226 SEKSENEKDLRKSELOQTALGNLKOIYYNSKAITSSSEKSADOFNTLLPKGFTG 285
DB 1 SEKSENEKDLRKSELOQTALGNLKOIYYNEKAKTENKESHDDQFQHTILFKGFTD 60
QY 286 HPWYNDLLVLGTAATSEYEGSSVDLYGAYGYQCAGGTGNKTCMYGGVTLHDNNRLT 345
DB 61 HSWYNDLLVDFDSKDIVDKYGGKVDLYGAYGYQCAGGTGNKTCMYGGVTLHDNNRLT 120
QY 346 EEKVPINLWDIGKQTTPIDKVTSSKEVTVQELDQARHYLHGKFGLYNSDSFGGKVQ 405
DB 121 EEKVPINLWDIGKQTTPIDKVTSSKEVTVQELDQARHYLHGKFGLYNSDSFGGKVQ 405

Db 121 EEKVPINLWDIGKQTTPIDKVTSSKEVTVQELDQARHYLHGKFGLYNSDSFGGKVQ 180
QY 406 RGLIVFHSSEGSTVSYDLFDAOGQYPTDLLRIYRDNTTISSTSLISLYLYTT 458
DB 181 RGLIVFHTSTEPSVNYDLFGAQQGQISNTLLRIYRDNKTNSNMHIDIYLYTS 233

RESULT 50

AAAB67338
ID AAB67338 standard; peptide; 233 AA.
XX
AC AAB67338;

DT 23-APR-2001 (first entry)
XX
DE Staphylococcus aureus enterotoxin A protein.
XX
KW Tumour; cancer; immune; enterotoxin.
XX
OS Staphylococcus aureus.

PN US6180097-B1.

PD 30-JAN-2001.

PF 30-OCT-1998; 98US-0183437.

PR 31-JAN-1994; 94US-0189424.

PR 19-JUN-1995; 95US-0491746.

PR 03-OCT-1989; 89US-0416530.

PR 17-JAN-1990; 90US-0466577.

PR 17-JAN-1991; 91WO-US00342.

PR 01-JUN-1992; 92US-0891718.

PR 02-MAR-1993; 93US-0025144.

XX (TERM/) TERMAN D S.

XX Terman DS;

XX WPI; 2001-1586657/16.

XX Tumor cell capable of stimulating antitumor immune reactivity in vitro

or in vivo comprises exogenous nucleic acids encoding a superantigen

and a costimulatory molecule

Disclosure; Fig 2; 16pp; English.

The present invention relates to a tumour cell capable of stimulating
antitumor immune reactivity in vitro or in vivo contains and
expresses an exogenous nucleic acid molecule encoding a superantigen
or its active fragment and an exogenous nucleic acid molecule
encoding a costimulatory molecule that activates T cells in
conjunction with an antigenic stimulus. The invention may be used
for cancer therapy by stimulating an anticancer immune response
in vivo or ex vivo.

Sequence 233 AA;

Query Match 26.7%; Score 941; DB 22; Length 233;
Best Local Similarity 76.0%; Pred. No. 5,1e-49;
Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

QY 226 SEKSENEKDLRKSELOQTALGNLKOIYYNSKAITSSSEKSADOFNTLLPKGFTG 285
DB 1 SEKSENEKDLRKSELOQTALGNLKOIYYNEKAKTENKESHDDQFQHTILFKGFTD 60
QY 286 HPWYNDLLVLGTAATSEYEGSSVDLYGAYGYQCAGGTGNKTCMYGGVTLHDNNRLT 345
DB 61 HSWYNDLLVDFDSKDIVDKYGGKVDLYGAYGYQCAGGTGNKTCMYGGVTLHDNNRLT 120

QY 346 EEKVPINLWDIGKQTTPIDKVTSSKEVTVQELDQARHYLHGKFGLYNSDSFGGKVQ 405
DB 121 EEKVPINLWDIGKQTTPIDKVTSSKEVTVQELDQARHYLHGKFGLYNSDSFGGKVQ 180

Qy 406 RGLIVPHSSEGSTVSYDLEDAQQYDPDTLLRIYRDNTTISLSISLYLYTT 458
| | | | | : | : | | | | | | | | | | | | | | | | : | : | | | :
Db 181 RGLIVPHTSTEPSVNYDLFGAQQYSNTLLRIYRONKTINSEMHIDIYLYTS 233

Search completed: January 8, 2003, 11:56:50
Job time : 61.5077 secs

THIS PAGE BLANK (USPTO)


```
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-013
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 122:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-360-107A-122

Query Match 31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKKSELOGTALGNLKOIYYNNKAITSEKXSADQFLNTLLFKGFFTG 285
Db 25 SEKSEINEKDLRKKSELORNALSRLQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
QY 286 HPWYNLLVDLGAATAATSEYSGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
Db 85 HPWYNLLVDLGSKDATNKYKGVLDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKVPINLWIDGKQTPVDIKVKTSSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 405
Db 145 EEKVPINLWIDGKQTPVDIKVKTSSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 204
QY 406 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLRYRDNNTTISSTLSLSLYTT 458
Db 205 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLRYRDNKNTINSENHLDLYTT 257

RESULT 4
US-08-484-223B-112
; Sequence 112, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
```

```
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-223B-112

Query Match 31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKKSELOGTALGNLKOIYYNNKAITSEKXSADQFLNTLLFKGFFTG 285
Db 25 SEKSEINEKDLRKKSELORNALSRLQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
QY 286 HPWYNLLVDLGAATAATSEYSGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
Db 85 HPWYNLLVDLGSKDATNKYKGVLDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKVPINLWIDGKQTPVDIKVKTSSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 405
Db 145 EEKVPINLWIDGKQTPVDIKVKTSSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 204
QY 406 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLRYRDNNTTISSTLSLSLYTT 458
Db 205 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLRYRDNKNTINSENHLDLYTT 257

RESULT 5
US-08-919-597-112
; Sequence 112, Application US/08919597
; Patent No. 6054265
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/919,597
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/470,896
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
```

REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-020
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-919-597-112

Query Match 31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;
QY 226 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFFTG 285
DB 25 SEKSEINEKDLRKSELORNALSRLQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
QY 286 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT 345
DB 85 HPWYNDLLVDLGSKDATNKYKGVLDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT 144
QY 346 EEKYPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 405
DB 145 EEKYPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 406 RGLIVFHSSEGSTSVSYDLFDAOGQYPTDLLRIYRDNNTTISSTSLSLYLTYT 458
DB 205 RGLIVFHSSEGSTSVSYDLFDAOGQYPTDLLRIYRDNNTTISSTSLSLYLTYT 257

RESULT 6
US-08-475-668A-112
Sequence 112, Application US/08475668A
Patent No 6060065
GENERAL INFORMATION:
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 211
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/475,668A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-026
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 112:

SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-475-668A-112
Query Match 31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;
QY 226 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFFTG 285
DB 25 SEKSEINEKDLRKSELORNALSRLQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
QY 286 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT 345
DB 85 HPWYNDLLVDLGSKDATNKYKGVLDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT 144
QY 346 EEKYPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 405
DB 145 EEKYPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 406 RGLIVFHSSEGSTSVSYDLFDAOGQYPTDLLRIYRDNNTTISSTSLSLYLTYT 458
DB 205 RGLIVFHSSEGSTSVSYDLFDAOGQYPTDLLRIYRDNNTTISSTSLSLYLTYT 257

RESULT 7
US-08-485-551A-112
Sequence 112, Application US/08485551A
Patent No 6068973
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
TITLE OF INVENTION: TRANSMISSION
NUMBER OF SEQUENCES: 211
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,551A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-023
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid

```

; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-551A-112

Query Match          31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLKKSELOGTALGNLKQIYYYNKAITSEKSAOFLTNWTLFLKGFFTG 285
      |||||
Db 25 SEKSEINEKDLKKSELORNALSNLRQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
      |||||

QY 286 HPWYNLLVDLGSTATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
      |||||
Db 85 HPWYNLLVDLGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
      |||||

QY 346 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 405
      |||||
Db 145 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 204
      |||||

QY 406 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLIYRDNTTISSTLSLSLYLYTT 458
      |||||
Db 205 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLIYRDNKTINSENLHIDLILYLYTT 257
      |||||

RESULT 8
US-08-471-913A-112
; Sequence 112, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/471,913A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
```

```

US-08-471-913A-112

Query Match          31.4%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 6.9e-67;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLKKSELOGTALGNLKQIYYYNKAITSEKSAOFLTNWTLFLKGFFTG 285
      |||||
Db 25 SEKSEINEKDLKKSELORNALSNLRQIYYNEKAITENKESDDQFLENTLLFKGFFTG 84
      |||||

QY 286 HPWYNLLVDLGSTATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
      |||||
Db 85 HPWYNLLVDLGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
      |||||

QY 346 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 405
      |||||
Db 145 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 204
      |||||

QY 406 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLIYRDNTTISSTLSLSLYLYTT 458
      |||||
Db 205 RGLIVFHSSEGSTSVSYDLFDAQGYPDTLRLIYRDNKTINSENLHIDLILYLYTT 257
      |||||

RESULT 9
US-08-485-264A-112
; Sequence 112, Application US/08485264A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/485,264A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-264A-112

Query Match          31.4%; Score 1107; DB 4; Length 257;
```

Best Local Similarity	89.7%;	Pred. No. 6.9e-67;			
Matches	209;	Conservative	9;	Mismatches	15;
				Indels	0;
Gaps	0;				
QY	226	SEKSEINEKDLRKXSELGQTALGNLKOIYYYNKAITSESSEKSAQOFLTNTLLFKGFFTG	285		
DB	25	SEKSEINEKDLRKXSELGQTALGNLKOIYYYNKAITSESSEKSAQOFLTNTLLFKGFFTG	84		
QY	286	HPWYNLLVLDGTAATSEYEGSSVDLLGAYGYQOCAGGTGPNKTACMYGGVTLHDNNRLT	345		
DB	85	HPWYNLLVLDGTAATSEYEGSSVDLLGAYGYQOCAGGTGPNKTACMYGGVTLHDNNRLT	144		
QY	346	EKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHLYHGKFGLYNSDSFGKVKQ	405		
DB	145	EKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHLYHGKFGLYNSDSFGKVKQ	204		
QY	406	RGLIVHSEGSTVSVDLFDAGQGYPTDLLRIYRDNNTTISLSLYLYTT	458		
DB	205	RGLIVHSEGSTVSVDLFDAGQGYPTDLLRIYRDNNTTISLSLYLYTT	257		
RESULT 10					
US-08-474-349A-112					
; Sequence 112, Application US/08474349A					
; Patent No. 6333395					
; GENERAL INFORMATION:					
; APPLICANT: Bolognesi, Dani P.					
; APPLICANT: Matthews, Thomas J.					
; APPLICANT: Wild, Carl T.					
; APPLICANT: Barney, Shawn O.					
; APPLICANT: Lambert, Dennis M.					
; APPLICANT: Petteway, Stephen R.					
; APPLICANT: Langlois, Alphonse J.					
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE					
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA					
; NUMBER OF INVENTION: VIRUS TRANSMISSION					
; NUMBER OF SEQUENCES: 517					
; CORRESPONDENCE ADDRESS:					
; ADDRESSEE: Pennie & Edmonds					
; STREET: 1155 Avenue of the Americas					
; CITY: New York					
; STATE: New York					
; COUNTRY: USA					
; ZIP: 10036-2711					
; COMPUTER READABLE FORM:					
; MEDIUM TYPE: Floppy disk					
; COMPUTER: IBM PC compatible					
; OPERATING SYSTEM: PC-DOS/MS-DOS					
; SOFTWARE: PatentIn Release #1.0, Version #1.30					
; CURRENT APPLICATION DATA:					
; APPLICATION NUMBER: US/08/474,349A					
; FILING DATE: 07-JUN-1995					
; CLASSIFICATION:					
; ATTORNEY/AGENT INFORMATION:					
; NAME: Coruzzi, Laura A.					
; REGISTRATION NUMBER: 30,742					
; REFERENCE/DOCKET NUMBER: 7872-024					
; TELECOMMUNICATION INFORMATION:					
; TELEPHONE: (212) 790-9090					
; TELEFAX: (212) 869-9741/8864					
; TELEX: 66141 PENNIE					
; INFORMATION FOR SEQ ID NO: 112:					
; SEQUENCE CHARACTERISTICS:					
; LENGTH: 257 amino acids					
; TYPE: amino acid					
; STRANDEDNESS:					
; TOPOLOGY: unknown					
; MOLECULE TYPE: protein					

Qy 300 AATSEYEGSSVDLYGAYYGOCAGGTPNKTACMYGGVTLHDNNRLTEBKKVPINLWIDGK 359
Db 254 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD-- 278
Qy 360 OTTVPIDKVKTSKEV-----TVQELDLQARHYLHCK-FGL-YNDSFSGKVGORGLIV 410
Db 279 --DVEHTAQTQPREEQNFSTRSVSELPIMHQDMLNKEFKRCRVNSAAFPAPIEK---- 332
Qy 411 FHSSEGSTSVSYDLFDAQGYQDPTLLRIYRDNTTISSTLSISLYLYTTSIVMTQPTSL 470
Db 333 -----TISK----- 337
Qy 471 VSAGDRVITCKASOSVNDVANYQQKPGQSPKLLISYSSRYAGVDPDRFSGSGYGTDF 530
Db 338 ----- 337
Qy 531 LTISVQAEDAANYFCOQDYNPPFTGGGKLEIKRADAAPTVSIFFPSSBOLTSGGASV 590
Db 338 -----KGRPKAPOVYTIPPKEQMAKDKVSL 363
Qy 591 VCFLNFPKIDINVKWKIDGSRQNGVLNSWTDODSKDSTYSMSSTLTLTDEYERHNSY 650
Db 364 TCMITDFPEDITVEMQWNGQPAEN-YKNTQPIMDT-DGSYFVYSKLVQKSNWEAGNTF 421
Qy 651 TCEATHK 657
Db 422 TCSVLHE 428

RESULT 12

US-08-353-400-36
; Sequence 36, Application US/08353400
; Patent No. 5665357
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 37
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/353.400
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9324819.3
; FILING DATE: 03-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9411089.7
; FILING DATE: 03-JUN-1994
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 464 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-353-400-36

Query Match 30.1%; Score 1059.5; DB 1; Length 464;
Best Local Similarity 37.5%; Pred. No. 2.2e-63;
Matches 250; Conservative 61; Mismatches 107; Indels 249; Gaps 14;

Qy 1 EVOLQSQPDLYKVCASVKISKASGYFTGYHHWVKQSGKGLWIGRINPNNGVTLY 60
Db 20 QVQLQQPQAEVLVPGASVOLSKASGYFTGYIHWVKQPGQGLWIGENPNSTGRSDY 79
Qy 61 NOKFKDKATLVKDSSTTAYMELASLTSEDSAVYYCARSTMT-TNYVMDYQGQTSVTVS 119
Db 80 NEKFNKATLVKDSSTTAYMQLSLTSEDSAVYYCARERAYGDAMDYQGQTSVTVS 139

Qy 120 SAKTTPPSVYPLAGSAAQTNMVMVTLGCLVKGYPEPVTVTWNSGSLSSGVHTFPVILQS 179
Db 140 SAKTTPPSVYPLAGSAAQTNMVMVTLGCLVKGYPEPVTVTWNSGSLSSGVHTFPVILQS 199
Qy 180 DLYTLSSSVTPSSVTPSETVTCNVAHPASSTKVDDKLVPRDSGPGSPSEKSEINEKDLRK 239
Db 200 DLYTLSSSVTPSSVTPSETVTCNVAHPASSTKVDDKLVPRDCG----- 243
Qy 240 KSELGQTALGNLKOIYYNYSKAITSSSEKSAQOFLTLLFKGFTGHWPYNDLLVDLGST 299
Db 244 -CKPCICIVPEVSSVFIPPK-----PKDVLITIL----- 272
Qy 300 AATSEYEGSSVDLYGAYYGOCAGGTPNKTACMYGGVTLHDNNRLTEBKKVPINLWIDGK 359
Db 273 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD-- 297
Qy 360 OTTVPIDKVKTSKEV-----TVQELDLQARHYLHCK-FGL-YNDSFSGKVGORGLIV 410
Db 298 --DVEHTAQTQPREEQNFSTRSVSELPIMHQDMLNKEFKRCRVNSAAFPAPIEK---- 351
Qy 411 FHSSEGSTSVSYDLFDAQGYQDPTLLRIYRDNTTISSTLSISLYLYTTSIVMTQPTSL 470
Db 352 -----TISK----- 356
Qy 471 VSAGDRVITCKASOSVNDVANYQQKPGQSPKLLISYSSRYAGVDPDRFSGSGYGTDF 530
Db 357 ----- 356
Qy 531 LTISVQAEDAANYFCOQDYNPPFTGGGKLEIKRADAAPTVSIFFPSSBOLTSGGASV 590
Db 357 -----KGRPKAPOVYTIPPKEQMAKDKVSL 382
Qy 591 VCFLNFPKIDINVKWKIDGSRQNGVLNSWTDODSKDSTYSMSSTLTLTDEYERHNSY 650
Db 383 TCMITDFPEDITVEMQWNGQPAEN-YKNTQPIMDT-DGSYFVYSKLVQKSNWEAGNTF 440
Qy 651 TCEATHK 657
Db 441 TCSVLHE 447

RESULT 13

US-08-896-933-24
; Sequence 24, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896.933
; CURRENT FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-24

Query Match 29.5%; Score 1040; DB 4; Length 226;
Best Local Similarity 85.2%; Pred. No. 1.8e-62;
Matches 196; Conservative 13; Mismatches 17; Indels 4; Gaps 1;

Qy 229 SEETNEKDLRKSELOQTALGNLKOIYYNYSKAITSSSEKSAQOFLTNTLLFKGFTGHWP 288
Db 1 SEETNEKDLRKSELOQTALGNLKOIYYNYSKAITSSSEKSAQOFLTNTLLFKGFTGHWP 60
Qy 289 YNDLLVDLGLSTAAATSEYEGSSVDLYGAYYGOCAGGTPNKTACMYGGVTLHDNNRLTEK 348
Db 61 YNDLLVDKSGDATNKYKGVLDLYGAYYGOCAGGTPNKTACMYGGVTLHDNNRLTEE- 119

Qy 349 KVPINLWIDGKQTTVPIDKVTSSKEVTQVLDLQARHLYHGKFGLYNSDSFGGKVQRL 408
Db 120 ---VBKWDGKQTTVPIDKVTSSKEVTQVLDLQARHLYHGKFGLYNSDSFGGKVQRL 176
Qy 409 IVFHSSEGSTVSYDLFDAQGYPTDQLRIYRDNNTTISSTLSISLYLTT 458
Db 177 IVFHSSEGSTVSYDLFDAQGYPTDQLRIYRDNNTTISSTLSISLYLTT 226

RESULT 14

US-09-314-235-24
; Sequence 24, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314.235
; CURRENT FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 08/896.933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252.978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891.718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466.577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416.530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 24
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-24

Query Match 29.5%; Score 1040; DB 4; Length 226;
Best Local Similarity 85.2%; Pred. No. 1.8e-62;
Matches 196; Conservative 13; Mismatches 17; Indels 4; Gaps 1;

Qy 229 SEETNEKDLRKKSELOGTALGNLKOIYYNSKAITSSSEKSADQELTNTLLKFGFTGHPW 288
Db 1 SEETNEKDLRKKSELOGTALGNLKOIYYNSKAITENKESDDQELNTLLKFGFTGHPW 60
Qy 289 YNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTTEK 348
Db 61 YNDLLVDLKGSKDATNKYKGVLDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEE- 119
Qy 349 KVPINLWIDGKQTTVPIDKVTSSKEVTQVLDLQARHLYHGKFGLYNSDSFGGKVQRL 408
Db 120 ---VBKWDGKQTTVPIDKVTSSKEVTQVLDLQARHLYHGKFGLYNSDSFGGKVQRL 176
Qy 409 IVFHSSEGSTVSYDLFDAQGYPTDQLRIYRDNNTTISSTLSISLYLTT 458
Db 177 IVFHSSEGSTVSYDLFDAQGYPTDQLRIYRDNNTTISSTLSISLYLTT 226

RESULT 15

US-09-485-737B-90
; Sequence 90, Application US/09485737B
; Patent No. 6350860
; GENERAL INFORMATION:
; APPLICANT: Buysse, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC SHOCK,
; FILE REFERENCE: INNS:015
; CURRENT APPLICATION NUMBER: US/09/485.737B

; CURRENT FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
US-09-485-737B-90

Query Match 27.8%; Score 980.5; DB 4; Length 711;
Best Local Similarity 35.3%; Pred. No. 7.6e-58;
Matches 254; Conservative 103; Mismatches 180; Indels 183; Gaps 23;

Qy 1 EVOIQSGPDLVKPGASVKISKASGYSTGYMHWKQSGKLEWIGRINPNNGVTLY 60
Db 21 QVQLVQSGSELKPGASVKISKASGYSTGYMHWKQAPQGLKMWGINTYTGSTY 80
Qy 61 NQKFKDQATLTVDKSSTTAYMELSLTSEDSAVYVCARSTMTITNYMDYWGOGTSVTYSS 120
Db 81 VDDFKGRVFSLDTSVSAAYLQISLKAEDATYFCARGF---YAMDYWGOGTIVTVSS 137
Qy 121 AKTTPPSVYPLAPGSAQTNSMTIGCLVKGYFPEPTVTWNSGSLSSGVHTFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAAALGCLVKDYFPEPTVSNNGALTSGVHTFPAVLQSS 197
Qy 181 -LYTLSSSVTPSPSTPSETVTCNVAHPASSTKYDKKIVPRDS-----GG 224
Db 198 GLYSLSVYVTPSSSLGTQTYICNVNHPKPSNTKVDKRVKPKSCDKTHTCPCPAPPELLGG 257
Qy 225 PS-----EKSEI-----NEKDLRKKSELOGTALGNLK---QIYYN 258
Db 258 PSVFLFPKPKDQTLMSRTPEVTCVVDVSHEDPEVFNWYDGVGVHNAKTKPREQYN 317
Qy 259 S-----KAITSS-EK-----SADQ 271
Db 318 STYRVSVSLTVLHQDWLNGKEYCKKVSNNKALPASIEKTSKAKGQPREPQVYVTLPPSREE 377
Qy 272 FLTN-----TLFKGFTGHPWYNDLLVDLGSTA-ATSEYEGSS--VDLYGAYGYQCAGG 324
Db 378 MTKNQVSLTCLVKGFYPP-----SDIAVEWESNGQPENNYKTTTPPVLDSDGSFFLYSLK-- 430
Qy 325 TPNKTAACHYGGV---TLHD--NNRLTEK-----KVPINLWIDGKQTTVPIDKV 368
Db 431 TVDKSRWQOGNVFSCSVNHEALHNYTOKLSLSPGKLGSGSQVQLVQSGSELKPKGASV 490
Qy 369 KTSKKEVTQVLDLQARHLYHGKFGLYNSDSFGGKVQRLIVFHSSEGSTVSYDLFDAQG 428
Db 491 KISCK-----ASGYTFTDYGMNWKQAPGQGLKMWGINTYTGSTYVD--DFKG 538
Qy 429 QYPTDQLRIYRDNNTTISSTLSIS-----LYLT----- 457
Db 539 RF-----VFLSDTSVSAAYLQISLKAEDATYFCARRGFYAMDYWGOGTIVTVSSGGG 592
Qy 458 -----TSIVMTQTPPTSLVSGADRVTTITCKASQSVSNDVAVYQOKPGSPKLLI 506
Db 593 GSGGGGSGGGSDIVLTQSPATMSASGGERVTLTCSASSISY-MFWYHQRPQSGSPRLII 651
Qy 507 SYTSSRYAGVPDRFSGGCGYGTFTLTISVQAEDAAYFCQQDYNSPPTFGGKTKLEIKR 566
Db 652 YDTSNLAGVPAFSGGSGGTYSLTISRMEPEDFATYFCHQSSSYPTFGGKTKLEIKR 711

RESULT 16
US-08-116-247-7
; Sequence 7, Application US/08116247

Query Match	27.4%	Score	966;	DB 2;	Length	468;			
Best Local Similarity	34.9%	Pred.	No. 4.3e-57;						
Matches	241;	Conservative	56;	Mismatches	136;	Indels	268;	Gaps	14;

QY	1	EVOLQQSGDPLVKPGASVKISCKASGYFTCYHHWVKQSPGKLEWIGTRINPNNGVLY	60
DB	20	QVQLQQSGAEELARPGASVKNSCKASGYFTFRYTHHWVKQRPGQGLEWIGLYINPSRGYTN	79
QY	61	NQPKFKKATITVQKSSPTAYMELRSLTSEDSAVYYCARSTMTINVMYDMYQGQTSVTYSS	120
DB	80	NQKFKDKATITLTDKSSSTAYMQLSSLTSEDSAVYYCAR--YYDDHYCLDYWGQGTTLTVSS	138
QY	121	AKTTPPSVYPLAPGSAQAQTNSVMTLGLCLVKGYFPEPVPTVTHNSGSLSSGVHTFAVLQSD	180
DB	139	AKTTPSVYPLAPVCGGDTTGSSVTLGLCLVKGYFPEPVTLTWNSSGSLSSGVHTFAVLQSD	198
QY	181	LYTLSSSVTPSPSTWPSSETVTCNVAHPASSTKVDPKIVPRDGGGSPSEKSEINEKDLRKK	240
DB	199	LYTLSSSVTPSTWPSQSITCNVAHPASSTKVDPKIEPR--GPTIKPCP-----	246
QY	241	SELQGTALGNLKQIYYNYSKAITSSSEKSAOFLTNTLLFKGFTTGHWPYNOLLVDLGSTA	300
DB	247	-----	246
QY	301	ATSEYEGSSVDLYGAYYGOCAGGTPNKTCMYGCVTLHDNNRLTTEKKVPINLWIDGKQ	360
DB	247	-----PCKCPAPN-----LLGGPSVF-----	262
QY	361	TTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVGORGLIVFHSSEGSTVS	420
DB	263	-----IFPPKIKDVLMI---SLSPITV	281

Db 2 ESGTELAKPGASVKMSCKASGYFTFSWIHWKQRPQGLEWIGYINPSTDYTYEIQKFK 61
Qy 66 DKATLTVDKSTTAYMELRSLTSDSAVYYCARSTMTITNYVMDYWGOGTSVTSSAKTTP 125
Db 62 DKATLTADKSSSTAYMQLSLSLTSVSSAKTTP 113
Qy 126 PSVYPLAPGSAOATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSDLYTLS 185
Db 114 PSVYPLAPGSAOATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSDLYTLS 173
Qy 186 SSVYTPSPSTWPEVTTCNVAHPASSTKVDKKIVPRD 221
Db 174 SSVYTPSPSTWPEVTTCNVAHPASSTKVDKKIVPRD 209

RESULT 18
US-08-303-569B-7
; Sequence 7, Application US/08303569B
; Patent No. 5859205
; GENERAL INFORMATION:
; APPLICANT: Adair, John R.
; APPLICANT: Athwal, Diljeet S.
; APPLICANT: Emage, John S.
; TITLE OF INVENTION: Humanised Antibodies
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5859205ris
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/303.569B
; FILING DATE: 07-SEP-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yako
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CARP-0032
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 468 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-303-569B-7

Query Match 27.2%; Score 957; DB 2; Length 468;
Best Local Similarity 34.7%; Pred. No. 1.6e-56;
Matches 240; Conservative 56; Mismatches 127; Indels 268; Gaps 14;
Qy 1 EVOLQSQGDLVKPGASVKISCKASGYFTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db 20 QVOLQSQGAELARPGASVKMSCKSGYFTTRYTMHWKQRPQGLEWIGYINPSRGYTN 79
Qy 61 NQKFKDKATLTVDKSTTAYMELRSLTSDSAVYYCARSTMTITNYVMDYWGOGTSVTSS 120
Db 80 NQKFKDKATLTVDKSTTAYMELRSLTSDSAVYYCARSTMTITNYVMDYWGOGTSVTSS 138
Qy 121 AKTTPSVYPLAPGSAOATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSD 180
Db 139 AKTTPSVYPLAPGCDTGTSSVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQSD 198
Qy 181 LYTSSSVTPSPSTWPEVTTCNVAHPASSTKVDKKIVPRDSDSGPSEKSEINEKDLRKK 240

Db 199 LYTSSSVTPSPSTWPEVTTCNVAHPASSTKVDKKIEPR---GPTIKPCP----- 246
Qy 241 SELQGTALGNLQIYYNSKRAITSSSEKSAQDLTNTLLFKGFFTGHPWYNDLLVDLAGSTA 300
Db 247 ----- 246
Qy 301 ATSEYEGSSVDLYGAYYGOCAGGTGPNKTCACMGVTLHDNNRLTEKKKVPINLWIDGKQ 360
Db 247 -----PCKCPAPN-----LLGGPSVF----- 262
Qy 361 TTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVGQRLIVPHSSSEGSTVS 420
Db 263 -----IPPKIKDVLMI---SLSPIVT 281
Qy 421 YDLFDAQGOYDPTLLRIYRDNMTTISLSTLSLYLTTTIVMTQTPTSLTLLVSAGDRVTIT 480
Db 282 CVVDVDSDDPD-----VOISWFEVNNVEVHTAQQT----- 312
Qy 481 CKASQSVSNDVAMYOQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTLTITSSVQAE 540
Db 313 -----HREDYNSTLRV----- 324
Qy 541 AAYVFCQDYNPPTFGGCTKLEIKRAD-----AAPTYSIFPPSSEOLT 584
Db 325 SALPIQHQDMWSGKEF-----KCKVNNKDLPAPIERTISKPGSVRAPQVYVLPPEEEMT 380
Qy 585 SGGASVVCFLANFYPKDINVKIKIDGSRQN-----GYLNSWTDQDSKDSYSSMSSTLT 639
Db 381 KKQVTLTCMTVDMPEDIYVETWNGKTELNYKNTEPVLD-----DGSYFMYSKLRV 433
Qy 640 TKDEYERHNSYTCEATHK-TSTSPIVKSFN 669
Db 434 EKKNNVERNYSYCSVWHEGLHNHHTTKSFSR 464

RESULT 19
PCT-US94-14106-51
; Sequence 51, Application PC/TUS9414106
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Process for Generating Specific Antibodies
; NUMBER OF SEQUENCES: 61
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14106
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 51:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-14106-51

Query Match 27.1%; Score 953.5; DB 5; Length 223;
Best Local Similarity 83.0%; Pred. No. 1.1e-56;
Matches 185; Conservative 12; Mismatches 23; Indels 3; Gaps 2;
Qy 1 EVOLQSQGDLVKPGASVKISCKASGYFTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db 1 EVOLQSQGPELMPGASVKISCKATGYTLSSYWLKQSPGCHGLEWIGELIFLGSGSAHY 60
Qy 61 NQKFKDKATLTVDKSTTAYMELRSLTSDSAVYYCARSTMTITNY--VMDYWGOGTSVT 118
Db 61 NEKFKGKATFTVDTSSNTAYMQLSLSLTSSEDSAVYYCARGD-YGNYGDYFDYWGQGTTLTV 119
Qy 119 SSAKTTTPSVYPLAPGSAOATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFFPAVLQ 178

Db 120 SSAKTPPSVYPLAPGSAQTNSMYTLCLLVKGYFPEPVTVTNWNSGLSSGSHVTFPAVLQ 179
Qy 179 SDLYTLSSSVTPSPSEVTVCNVAHPASSTKVKDKKIVPRD 221
Db 180 SDLYTLSSSVTPSPSEVTVCNVAHPASSTKVKDKKIVPRD 222

RESULT 20
US-08-486-099-113
: Sequence 113, Application US/08486099
: Patent No. 6013263
: GENERAL INFORMATION:
: APPLICANT: Bolognesi, Dani P.
: APPLICANT: Matthews, Thomas J.
: APPLICANT: Wild, Carl T.
: APPLICANT: Barney, Shawn O.
: APPLICANT: Lambert, Dennis M.
: APPLICANT: Petteway, Stephen R.
: APPLICANT: Langlois, Alphonse J.
: TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
: TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
: TITLE OF INVENTION: B VIRUS TRANSMISSION
: NUMBER OF SEQUENCES: 209
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/486,099
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7872-031
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 113:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 257 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-486-099-113

Query Match 26.9%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 226 SEKSEINEKDLRKSELGQTALGNLKOIYYNYSKAITSSSEKSADQFLTNTLLFKGFFTG 285
Db 25 SEKSEINEKDLRKSELGQTALGNLKOIYYNYSKAITSSSEKSADQFLTNTLLFKGFFTG 84

Qy 286 HPWNLDLVLGSTAAATSEYEGSSVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLT 345
Db 85 HSWNDLLVDFDSDIVDKYKGVKVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLT 144

Qy 346 EEKVPINLWIDGKQTTVPIDKVKTSKEVTQVELDQARHYLHGKFGLYNSDSFGGKVQ 405
Db 145 EEKVPINLWIDGKQNTVPLETVTKNKNVTQVELDQARRYLQEKYNLYNSDVFDDGVQ 204

Qy 406 RGLIVFHSSEGSTVSYDLFDAQQGYQPTLLRIYRDNTTISSTLSISLYLTT 458
Db 205 RGLIVFHTSTEPSVNYDLFGAQQGYSNLTLRIYRDNKTINSENHIDIYLYTS 257

RESULT 21
US-08-360-107A-123
: Sequence 123, Application US/08360107A
: Patent No. 6017536
: GENERAL INFORMATION:
: APPLICANT: Bolognesi, Dani P.
: APPLICANT: Matthews, Thomas J.
: APPLICANT: Wild, Carl T.
: APPLICANT: Barney, Shawn O.
: APPLICANT: Lambert, Dennis M.
: APPLICANT: Petteway, Stephen R.
: APPLICANT: Langlois, Alphonse J.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
: TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
: TITLE OF INVENTION: TRANSMISSION
: NUMBER OF SEQUENCES: 149
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/360,107A
: FILING DATE: 20-DEC-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7872-013
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 123:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 257 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-360-107A-123

Query Match 26.9%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 226 SEKSEINEKDLRKSELGQTALGNLKOIYYNYSKAITSSSEKSADQFLTNTLLFKGFFTG 285
Db 25 SEKSEINEKDLRKSELGQTALGNLKOIYYNYSKAITSSSEKSADQFLTNTLLFKGFFTG 84

Qy 286 HPWNLDLVLGSTAAATSEYEGSSVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLT 345
Db 85 HSWNDLLVDFDSDIVDKYKGVKVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLT 144

Qy 346 EEKVPINLWIDGKQTTVPIDKVKTSKEVTQVELDQARHYLHGKFGLYNSDSFGGKVQ 405
Db 145 EEKVPINLWIDGKQNTVPLETVTKNKNVTQVELDQARRYLQEKYNLYNSDVFDDGVQ 204

Qy 406 RGLIVFHSSEGSTVSYDLFDAQQGYQPTLLRIYRDNTTISSTLSISLYLTT 458
Db 205 RGLIVFHTSTEPSVNYDLFGAQQGYSNLTLRIYRDNKTINSENHIDIYLYTS 257

RESULT 22

US-08-484-223B-113
; Sequence 113, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-223B-113

Query Match 26.9%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy	226	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSSEKSADQFLTNTLLFKGFFTG	285
Db	25	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSSEKSADQFLTNTLLFKGFFTG	84
Qy	286	HPWYNDLLVGLGTAATSEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT	345
Db	85	HSWYNDLLVDFDSKDIVDKYKGVLDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT	144
Qy	346	EEKKVPINLWIDGKQTTPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDFGKQV	405
Db	145	EEKKVPINLWIDGKQTTPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDFGKQV	204
Qy	406	RGLIVFHSSEGSTVSVDLFDAGQGVPTDRLRIYRDNNTTISSTSLISLYLYTT	458
Db	205	RGLIVFHTSTEPSVNDLFDAGQGVPTDRLRIYRDNNTTISSTSLISLYLYTT	257

RESULT 23

US-08-919-597-113
; Sequence 113, Application US/08919597
; Patent No. 6054265
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/470,896
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-919-597-113

Query Match 26.9%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy	226	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSSEKSADQFLTNTLLFKGFFTG	285
Db	25	SEKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSSEKSADQFLTNTLLFKGFFTG	84
Qy	286	HPWYNDLLVGLGTAATSEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT	345
Db	85	HSWYNDLLVDFDSKDIVDKYKGVLDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT	144
Qy	346	EEKKVPINLWIDGKQTTPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDFGKQV	405
Db	145	EEKKVPINLWIDGKQTTPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDFGKQV	204
Qy	406	RGLIVFHSSEGSTVSVDLFDAGQGVPTDRLRIYRDNNTTISSTSLISLYLYTT	458
Db	205	RGLIVFHTSTEPSVNDLFDAGQGVPTDRLRIYRDNNTTISSTSLISLYLYTT	257

RESULT 24


```
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue Of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,264A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-485-264A-113

Query Match          26.9%; Score 948; DB 4; Length 257;
Best Local Similarity 76.4%; Pred. No.3e-56;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy    226 SEKSEINEKDLRKKSSEQALGNLKOIIYYNSKAITSSESQAQQFTLTNTLLFKGFFTG 285
      ||| | | | | | | | | | | | | | | | : | : | : | | | | | | | | | | |
Db    25   SEKSEINEKDLRKKSSEQALGNLKOIIYYNEAKTENKESHDDQLFQHITILFGGFETD 84
      ||| | | | | | | | | | | | | | | | : | : | : | | | | | | | | | | |

Qy    286 HPWYNLDLVLGSTAATSEYGSSVDLYGYGVOCAGTPNKTKACMYGGVTLHNNRLT 345
      | | | | | | | | | : | : | | | | | | | | | | | | | | | | | | | | |
Db    85 HSWYNLDLVDFEDSKVIDKYGKKVDDLGYAYIGYCAGTPNKTKACMYGGVTLHNNRLT 144
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy    346 EEKVVPINLWDGKTPTPIDBKVTTSKEEVTVEQLDQARHYLHGKFGLYSNDSFSGKVQ 405
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db    145 EEKVVPINLWDGQTNPTEVTNKNVVVELDQARRYLQEKEYNLNYNSDVFDGKVQ 204
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

Qy    406 RGLIVFHSSSEGSTVSVDLFDAQGPDPDLLRIYRDNTTTISLSISLYYTTF 458
      | | | | | : | : | | | | | | | | | | | | | | | | | : | : | : | : |
Db    205 RGLIVFHTSPSVNDLFGAQGSNTLLRIYRNKTI NSENMHIDI LYL TS     257


RESULT 28
US-08-474-349A-113
; Sequence 113, Application US/08474349A
; Patent No. 6333395
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
; TITLE OF INVENTION: VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 517
; CORRESPONDENCE ADDRESS:
```


ADDRESSEE: Pennle & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA: US/08/474,349A
APPLICATION NUMBER: US/08/474,349A
FILING DATE: 07-JUN-1995
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-024
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-474-349A-113

Query Match 26.9%; Score 948; DB 4; Length 257;

Best Local Similarity 76.4%; Pred. No. 3e-56;

Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSDAQFLTNTLLFKGFFTG 285
DB 25 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSDAQFLTNTLLFKGFFTD 84
QY 266 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGTGNKTACMYGGVTLHDNNRLT 345
DB 85 HSWYNDLLVDFDSDIVDKYKGVLDLYGAYGYOCAGTGNKTACMYGGVTLHDNNRLT 144
QY 346 BEKKVPINLWDGKQTTVPIDKVKTSKREVTQVELDLQARHYLHGKFLGFLYNSDSFGGRVQ 405
DB 145 BEKKVPINLWDGKQNTVPLETVTKNKNVTQVELDLQARRYLQEKYLNLYNSDVFGRVQ 204
QY 406 RGLIVFHSSEGSTVSYDLFDAGQGPDPDLLRIYRDNNTTISLSLSISLYLTT 458
DB 205 RGLIVFHTSTEPSVNYDLFGAQGOYSNTLLRIYRDNKTINSENMHIDIYLYTS 257

RESULT 29

US-08-353-400-37
Sequence 37, Application US/08353400
Patent No. 5665357
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 37
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/353.400
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9324819.3
FILING DATE: 03-DEC-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9411089.7
FILING DATE: 03-JUN-1994
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 239 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-353-400-37
Query Match 26.8%; Score 944.5; DB 1; Length 239;
Best Local Similarity 78.1%; Pred. No. 4.8e-56;
Matches 185; Conservative 22; Mismatches 19; Indels 11; Gaps 3;
QY 445 SSTSLISLSLYLTT----SIVMTQTPTSLLYSAGDRVTITCKASOSVND-----VAVY 494
DB 3 SOAQVLILLLLWVGTCGDIVMSQSPSSSLAVSAGEKVTMSCKSSQSLNSRTRKNYLAWY 62
QY 495 OOKPGOSPCKLLTSYTSRYAGVDPDRFSGGYGTDTFTLTISSVQAEADAAYVFCOODYNSPP 554
DB 63 QORPGOSPCKLLLYWASTRTSGVDPDRFTGSGGTDFTLTISSVQAEADLAIYCKQSY-TLR 121
QY 555 TFGGGTKLEIKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSRQ 614
DB 122 TFGGGTKLEIKRADAAPTVSIFPPSSEQLTSGGASVVCFLNNFYPKDINVKWKIDGSRQ 181
QY 615 NGVLNSWTDQDSKDSYMSSTLTLTCKDEYERHNSYTCETHKTSTSPIVKSFNRNE 671
DB 182 NGVLNSWTDQDSKDSYMSSTLTLTCKDEYERHNSYTCETHKTSTSPIVKSFNRNE 238

RESULT 30

US-08-446-918A-4
Sequence 4, Application US/08446918A
Patent No. 5705151
GENERAL INFORMATION:
APPLICANT: Dow, Steve W.
APPLICANT: Elmslie, Robyn E.
TITLE OF INVENTION: GENE THERAPY FOR T CELL REGULATION
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446.918A
FILING DATE: 18-MAY-1995
CLASSIFICATION: 552
ATTORNEY/AGENT INFORMATION:
NAME: Kovarik, Joseph E.
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2879-29
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-918A-4

Query Match 26.8%; Score 944; DB 1; Length 233;

Best Local Similarity 76.3%; Pred. No. 5e-56;

Matches 177; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 227 EKSEINEKDLRKSELQGTALGNLKIYYNKAITSSEKSAQOFLNTLLFKGFFTGH 286

Db 2 EKSEINEKDLRKSELQGTALGNLKIYYNKAITSSEKSAQOFLNTLLFKGFFTGH 61

Qy 287 PHYNLLVDLGSTAASTSEYEGSSVDLYGAYGYQCAGGTPNKTACMGYGVTLHDNNRLTE 346

Db 62 SWYNLLVDFDSKDIDVYKGGKVDLYGAYGYQCAGGTPNKTACMGYGVTLHDNNRLTE 121

Qy 347 EKKVPINLWIDGKQTPIDVKVTSKKEVTQVELDQARHYLHGKFGLYNSDSFGGKQVOR 406

Db 122 EKKVPINLWIDGKQTPIDVKVTSKKEVTQVELDQARHYLHGKFGLYNSDSFGGKQVOR 181

Qy 407 GLIVFHSSEGSTSVSYDLFDAGQOYPTDILLRIYRNTTISSTLSISLYLYTT 458

Db 182 GLIVFHTSTEPSVNDLFGAOGQYSNTLLRIYRNTTINSENMHIDILYLYTS 233

RESULT 31

US-08-580-806-4

; Sequence 4, Application US/08580806

; Patent No. 5935568

; GENERAL INFORMATION:

; APPLICANT: Dow, Steve W.

; APPLICANT: Elmslie, Robyn E.

; APPLICANT: Potter, Terence A.

; TITLE OF INVENTION: GENE THERAPY FOR EFFECTOR CELL REGULATION

; NUMBER OF SEQUENCES: 13

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sheridan Ross & McIntosh

; STREET: 1700 Lincoln Street, Suite 3500

; CITY: Denver

; STATE: Colorado

; COUNTRY: U.S.A.

; ZIP: 80203

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patentin Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/580.806

; FILING DATE:

; CLASSIFICATION: 424

; ATTORNEY/AGENT INFORMATION:

; NAME: Connell, Gary J.

; REGISTRATION NUMBER: 32,020

; REFERENCE/DOCKET NUMBER: 2879-29-C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (303) 863-9700

; TELEFAX: (303) 863-0223

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 233 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-580-806-4

Query Match 26.8%; Score 944; DB 2; Length 233;

Best Local Similarity 76.3%; Pred. No. 5e-56;

Matches 177; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 227 EKSEINEKDLRKSELQGTALGNLKIYYNKAITSSEKSAQOFLNTLLFKGFFTGH 286

Db 2 EKSEINEKDLRKSELQGTALGNLKIYYNKAITSSEKSAQOFLNTLLFKGFFTGH 61

Qy 287 PHYNLLVDLGSTAASTSEYEGSSVDLYGAYGYQCAGGTPNKTACMGYGVTLHDNNRLTE 346

Db 62 SWYNLLVDFDSKDIDVYKGGKVDLYGAYGYQCAGGTPNKTACMGYGVTLHDNNRLTE 121

ADDRESSEE: Wenderoth, Lind & Ponack
STREET: 805 Fifteenth Street, N.W., #700
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: MS-DOS
SOFTWARE: Wordperfect 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737.129A
FILING DATE: No. 5885816ember 15, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Warren M. Cheek, Jr.
REGISTRATION NUMBER: 33,367
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-8850
TELEFAX:
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 222 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-737-129A-6

Query Match 26.5%; Score 933; DB 2; Length 222;
Best Local Similarity 82.6%; Pred. No. 2.5e-55;
Matches 180; Conservative 12; Mismatches 24; Indels 2; Gaps 1;
QY 6 QSGPDLVKGASVVKISCKASGYFTGYVMHWVKQSPGKLEWIGRINPNNGVTLYNOKFK 65
DB 2 ESGPELVKGGVSTVVKISCKASGYFTTSMNVRQRPQCGLEWIGRIYPOSGDNNGYNGRKF 61
QY 66 DKATLVDSKSTAYMELRSLTSEDSAVYICAR--STMITNTYMDYWGOGTSVTYSSAKT 123
DB 62 VKATLTAERSSTTVVHLHLSLTSVDSAVYECARFHYDYRRSYAMDYWGOGTSVTYSSAKT 121
QY 124 TTPSVYPLAGSAAQTNSHWTLGCLVKGYPPEPVTVTNWNSGSLSGSVHTFPVAVLOS DLYT 183
DB 122 TTPSVYPLAGSAAQTNSHWTLGCLVKGYPPEPVTVTNWNSGSLSGSVHTFPVAVLOS DLYT 181
QY 184 LSSSVTPSPSTWPSSETVTCNVAPASSTKVDKKIVPRD 221
DB 182 LSSSVTPSPSTWPSSETVTCNVAPASSTKVDKKIVPRD 219

RESULT 34
US-08-896-933-23
Sequence 23, Application US/08896933
Patent No. 6221351
GENERAL INFORMATION:
APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005002
CURRENT APPLICATION NUMBER: US/08/896,933
CURRENT FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252,978
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 232

TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-896-933-23
Query Match 26.4%; Score 931.5; DB 4; Length 232;
Best Local Similarity 76.0%; Pred. No. 3.4e-55;
Matches 177; Conservative 21; Mismatches 34; Indels 1; Gaps 1;
QY 226 SEKSEINEKDLRKSELOGTALGNLKOIYYNNKAITSSSEKSAOQFLTNTLLFKGFFTG 285
DB 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNNKAITSSSEKSAOQFL-HTILFKGFFTD 59
QY 286 HPWYNDLLVDLSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHNNRLT 345
DB 60 HSWYNDLLVDFDSKDIDVKYKGGKVDLYGAYGYOCAGGTPNKTACMYGGVTLHNNRLT 119
QY 346 EEKVPINLWIDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQ 405
DB 120 EEKVPINLWIDGKQTTVPLETVTNKNVTVOELDPQARRYLQEKYLNLYNSDVFQKVKQ 179
QY 406 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTLSLSLYLYTT 458
DB 180 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTLSLSLYLYTT 232

RESULT 35
US-09-314-235-23
Sequence 23, Application US/09314235
Patent No. 6338845
GENERAL INFORMATION:
APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005004
CURRENT APPLICATION NUMBER: US/09/314,235
CURRENT FILING DATE: 1999-05-18
EARLIER APPLICATION NUMBER: 08/896,933
EARLIER FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252,978
EARLIER FILING DATE: 1994-06-02
EARLIER APPLICATION NUMBER: 07/891,718
EARLIER FILING DATE: 1992-06-01
EARLIER APPLICATION NUMBER: US91/00342
EARLIER FILING DATE: 1991-01-17
EARLIER APPLICATION NUMBER: 07/466,577
EARLIER FILING DATE: 1990-01-17
EARLIER APPLICATION NUMBER: 07/416,530
EARLIER FILING DATE: 1989-10-03
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 232
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-314-235-23

Query Match 26.4%; Score 931.5; DB 4; Length 232;
Best Local Similarity 76.0%; Pred. No. 3.4e-55;
Matches 177; Conservative 21; Mismatches 34; Indels 1; Gaps 1;
QY 226 SEKSEINEKDLRKSELOGTALGNLKOIYYNNKAITSSSEKSAOQFLTNTLLFKGFFTG 285
DB 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNNKAITSSSEKSAOQFL-HTILFKGFFTD 59
QY 286 HPWYNDLLVDLSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHNNRLT 345
DB 60 HSWYNDLLVDFDSKDIDVKYKGGKVDLYGAYGYOCAGGTPNKTACMYGGVTLHNNRLT 119
QY 346 EEKVPINLWIDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQ 405
DB 120 EEKVPINLWIDGKQTTVPLETVTNKNVTVOELDPQARRYLQEKYLNLYNSDVFQKVKQ 179
QY 406 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTLSLSLYLYTT 458

Db 180 RGLIVFHTSTPSVNDLFGAQQYQNTLLRIYRDNKTINSENHDIYLYTS 232

RESULT 36
US-09-144-776B-2
; Sequence 2, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
; Mark A. Olson
; Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
; Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army MRC -504 Scott Street
; MCMR-JA (Charles H. Harris-Patent
; Atty)
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,776B
; FILING DATE: 01-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/882,431
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-144-776B-2
Query Match 26.3%; Score 925; DB 4; Length 257;
Best Local Similarity 75.1%; Pred. No. 1.le-54;
Matches 175; Conservative 21; Mismatches 37; Indels 0; Gaps 0;
QY 226 SEKSEINEKDLRKSELOGTALGNLQIYYNKAITSSEKSADQFLNTLLFGFFTG 285
Db 25 SEKSEINEKDLRKSELOGTALGNLQIYYNKAITSSEKSADQFLNTLLFGFFTG 84
QY 286 HPWYNDLLVDLSTAASTSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
Db 85 HSWYNDLLVRFDSKDIDYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 346 EEKVPINLWIDGKQTPIDVKYKSKKVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 405
Db 145 EEKVPINLWIDGKQTPIDVKYKSKKVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 204
QY 406 RGLIVFHSSEGSTSVYDLFDAQQYQPDTLRLIYRDNKTINSENHDIYLYTT 458
Db 205 RGLIVFHTSTPSVNDLFGAQQYQNTLLRIYRDNKTINSENHDIYLYTS 257

RESULT 37
US-09-144-776B-4
; Sequence 4, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
; Mark A. Olson
; Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
; Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army MRC -504 Scott Street
; MCMR-JA (Charles H. Harris-Patent
; Atty)
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,776B
; FILING DATE: 01-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/882,431
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 233
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-144-776B-4

Query Match 26.1%; Score 921; DB 4; Length 233;
Best Local Similarity 75.0%; Pred. No. 1.7e-54;
Matches 174; Conservative 21; Mismatches 37; Indels 0; Gaps 0;
QY 227 EKSEINEKDLRKSELOGTALGNLQIYYNKAITSSEKSADQFLNTLLFGFFTG 286
Db 2 EKSEINEKDLRKSELOGTALGNLQIYYNKAITSSEKSADQFLNTLLFGFFTG 61
QY 287 HPWYNDLLVDLSTAASTSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 346
Db 62 HSWYNDLLVRFDSKDIDYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 121
QY 347 EEKVPINLWIDGKQTPIDVKYKSKKVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 406
Db 122 EEKVPINLWIDGKQTPIDVKYKSKKVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 181
QY 407 RGLIVFHSSEGSTSVYDLFDAQQYQPDTLRLIYRDNKTINSENHDIYLYTT 458
Db 182 RGLIVFHTSTPSVNDLFGAQQYQNTLLRIYRDNKTINSENHDIYLYTS 233
RESULT 38
US-08-792-824-10

```
; Sequence 10, Application US/08792824
; Patent No. 5932449
; GENERAL INFORMATION:
; APPLICANT: EMANUEL, PETER A.
; APPLICANT: BURANS, JAMES P.
; APPLICANT: VALDES, JAMES J.
; APPLICANT: MOHVEE, ELDEFRAWI E.
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. Army Chemical and Biological Defense
; STREET: Office of the Chief Counsel, Bldg E4435
; CITY: Aberdeen Proving Ground
; STATE: MD
; COUNTRY: U.S.
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08792.824
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Biffoni, U. J.
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 431-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158
; TELEFAX: 410-671-2534
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-792-824-10

Query Match 26.0%; Score 915.5; DB 2; Length 254;
Best Local Similarity 79.2%; Pred. No. 4.5e-54;
Matches 179; Conservative 15; Mismatches 27; Indels 5; Gaps 2;

Qy 1 EVLOQSGPDLVKPGASVKISKASGYSTGYMHVWKQSPKGLGWIGRINPNNGVTLY 60
Db 23 EVLOQSGAELVKPGASVKLSCTASGFNIKDTFHMVKQRPQGLEWIGRIDPANGNTEY 82
Qy 61 NOKFKDKATLTVDKSSTTAYMELRSLTSEDSAVYYCARSTMTITNYMDYWGOGTSVTYSS 120
Db 83 DPKFGKATITADTSSNTVNLQSLTSEDATAYYCASGGELG---FPYWGQGLTVTVSA 139
Qy 121 AKTTSPSYVPLAPGSAAGTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
Db 140 AKTTSPSYVPLAPGSAAGTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 199
Qy 181 LYTLSSTVPSSTWPSSETVTCNVAHPASSTKVDKKIIPRD---SGG 224
Db 200 LYTLSSTVPSSTWPSSETVTCNVAHPASSTKVDKKIIPRDCTSGG 245

RESULT 39
US-08-792-824-13
; Sequence 13, Application US/08792824
; Patent No. 5932449
; GENERAL INFORMATION:
; APPLICANT: EMANUEL, PETER A.
; APPLICANT: BURANS, JAMES P.
; APPLICANT: VALDES, JAMES J.
; APPLICANT: MOHVEE, ELDEFRAWI E.
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN
; NUMBER OF SEQUENCES: 13
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. Army Chemical and Biological Defense
; STREET: Office of the Chief Counsel, Bldg E4435
; CITY: Aberdeen Proving Ground
; STATE: MD
; COUNTRY: U.S.
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08792.824
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Biffoni, U. J.
; REGISTRATION NUMBER: 39,908
; REFERENCE/DOCKET NUMBER: DAM 431-96
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 410-671-1158
; TELEFAX: 410-671-2534
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 254 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-792-824-13

Query Match 26.0%; Score 915.5; DB 2; Length 254;
Best Local Similarity 79.2%; Pred. No. 4.5e-54;
Matches 179; Conservative 15; Mismatches 27; Indels 5; Gaps 2;

Qy 1 EVLOQSGPDLVKPGASVKISKASGYSTGYMHVWKQSPKGLGWIGRINPNNGVTLY 60
Db 23 EVLOQSGAELVKPGASVKLSCTASGFNIKDTFHMVKQRPQGLEWIGRIDPANGNTEY 82
Qy 61 NOKFKDKATLTVDKSSTTAYMELRSLTSEDSAVYYCARSTMTITNYMDYWGOGTSVTYSS 120
Db 83 DPKFGKATITADTSSNTVNLQSLTSEDATAYYCASGGELG---FPYWGQGLTVTVSA 139
Qy 121 AKTTSPSYVPLAPGSAAGTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
Db 140 AKTTSPSYVPLAPGSAAGTNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 199
Qy 181 LYTLSSTVPSSTWPSSETVTCNVAHPASSTKVDKKIIPRD---SGG 224
Db 200 LYTLSSTVPSSTWPSSETVTCNVAHPASSTKVDKKIIPRDCTSGG 245

RESULT 40
US-08-792-824-4
; Sequence 4, Application US/08792824
; Patent No. 5932449
; GENERAL INFORMATION:
; APPLICANT: EMANUEL, PETER A.
; APPLICANT: BURANS, JAMES P.
; APPLICANT: VALDES, JAMES J.
; APPLICANT: MOHVEE, ELDEFRAWI E.
; TITLE OF INVENTION: DETECTION OF BOTULINUM TOXIN
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: U.S. Army Chemical and Biological Defense
; STREET: Office of the Chief Counsel, Bldg E4435
; CITY: Aberdeen Proving Ground
; STATE: MD
; COUNTRY: U.S.
; ZIP: 21010-5423
; COMPUTER READABLE FORM:
```


Db	138	TVSIFPPSPSEQLTSGGASVVCFLNFPKDLNVKWKIDGSRQGVNLNSWTDQDQKDYTY	197
Qy	632	SMSSTLTLTKDYEYRHNSYTCATHKTKSTSPIVKSEFNRE	671
Db	198	SMSSTLTLTKDYEYRHNSYTCATHKTKSTSPIVKSEFNRE	237
RESULT 43			
	5455030-1		
	Patent No. 5455030		
	APPLICANT: LADNER, ROBERT C.;BIRD, ROBERT E.;HARDMAN, KARL		
	TITLE OF INVENTION: IMMUNOTHERAPY USING SINGLE CHAIN		
	POLYPEPTIDE BINDING MOLECULES		
	CURRENT APPLICATION DATA:		
	NUMBER OF SEQUENCES: 24		
	APPLICATION NUMBER: US/08/40,440		
	FILING DATE: 1-APR-1993		
	PRIOR APPLICATION DATA:		
	APPLICATION NUMBER: 512,910		
	FILING DATE: 25-APR-1990		
	APPLICATION NUMBER: 299,617		
	FILING DATE: 19-JAN-1989		
	APPLICATION NUMBER: 92,110		
	FILING DATE: 02-SEP-1987		
	APPLICATION NUMBER: 902,971		
	FILING DATE: 01-SEP-1986		
	SEQ ID NO:1:		
	LENGTH: 447		
	5455030-1		
	Query Match	25.7%;	Score 903.5; DB 6; Length 447;
	Best Local Similarity	32.9%;	Pred. No. 5.9e-53;
	Matches	222; Conservative	69; Mismatches 121; Indels 263; Gaps 15;
Qy	1	EVQLQQSPDLVKPGASVKISCKASGYSTFYHHWYKQSPGKGLEWIGRINPNNGVTLY	60
Db	1	EVQLVESGGDLVKKPG-SLKLSAASGFTFISYGSWVRQTPDKRLWATVSSGVTYY	59
Qy	61	NQKFKDKATLTVDKSSITAYNELKSLTSEDSAVYYCAR---STMITN-VYMDYWGQGTSV	116
		: : : : : : : : : : : : : : : : : : : : : : :	
Db	60	PDSVKGRFTISRDAKNTLYLQMSGLKSEDTAMYYCARRITTVLTLYANDYWGQGTSV	119
Qy	117	TVSSAKTTPPSVYPLAPGSAATNSMTVLGCLVKGYFPEPVTVTWNSGSLSSGVTTPAV	176
Db	120	TVSSAKTTPPSVYPLAPGSAATNSMTVLGCLVKGYFPEPVTVTWNSGSLSSGVTTPAV	179
Qy	177	LQSDLYLTSSSVTPSPSTWSPSETVTCNVAHPASTKVDKKIIPRDSGGPSEKSEINEKD	236
Db	180	LQSDLYLTSSSVTPSPSTWSPSETVTCNVAHPASTKVDKKIIPRDCG-----	226
Qy	237	LRKKSLEQGTALGNLKQIYYNNSKAITSSSEKSAQDLTNTLLEKGFTHGHPWNDLLVDL	296
		: : : : : : : : : : : : : : : : : : : : : : : :	
Db	227	---CKPCICTVPEVSVFIPPK-----PKDVLITL-----	255
Qy	297	GSTAATSEYGGSDLYGAYYGOCAGTTPNKTACMYGGVTLHDNNRLTEEKVPIINLWI	356
		: : : : : : : : : : : : : : : : : : : :	
Db	256	-----TP-KVTCVVVDIS-----KDDPEVQFSWEV	279
Qy	357	DGKQTTVPIDKVKTSKKEV-----TVQELDLQARHYLHCKFGLYINSDS--FGCKVQRG	407
Db	280	D-----DVEVHTAQTPREEQPDTSRSVSELPIMHQDMLNKEFKCRVDSSAAFPAPTEK-	334
Qy	408	LIVPHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNTTISSTSLISLYLYTTSIVMTQPTT	467
Db	335	-----TISKT-----	339
Qy	468	SLVVSAGDRVITITCKASQSVSNDAVYQKPGQSPKLLISYTSRYAGVPDRFSGSGYGT	527
Db	340	-----	339
Qy	528	DFTLTISVQNAEDAAVYFCQODYNSPPTFGGTTKLETKRADAAPTVSIFPPSSBOLTSFGG	587
Db	340	-----KGRKAPQVYTIPTPPKKQMAKDK	362

Qy	588	ASVVCFLNNFYPKDINVKWKIDGSRQN-----GVLSNWTDDQSKDGSTYSMSSTLTLLTKD	642
Db	363	VSLTMITDFEPEDITVEWQMDGPAENYKNTQRIMNT-----DGSYFVYSKLDVQKS	415
Qy	643	EYERHNSYTCEATHK	657
Db	416	NWEAGDTFTCSVLHE	430
RESULT 44			
PCT-US94-14106-57			
; Sequence 57, Application PC/TUS9414106			
; GENERAL INFORMATION:			
; APPLICANT:			
; TITLE OF INVENTION: Process for Generating Specific Antibodies			
; NUMBER OF SEQUENCES: 61			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: ASCII (text)			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: PCT/US94/14106			
; FILING DATE:			
; CLASSIFICATION:			
; INFORMATION FOR SEQ ID NO: 57:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 218 amino acids			
; TYPE: amino acid			
; TOPOLOGY: linear			
; MOLECULE TYPE: protein			
PCT-US94-14106-57			
Query Match 25.6%; Score 901; DB 5; Length 218;			
Best Local Similarity 80.6%; Pred. No. 3.5e-53;			
Matches 174; Conservative 9; Mismatches 29; Indels 4; Gaps			
Qy	460	IVMTQTPTSLLVASGDRVTTCRASQSVND----VANVQQKPGQSPKKLIISYTSRRYAG	515
Db	2	IVMTQSPASLAVSLGQRATISCRVRSVSSSHYMHVYQQKPGQPPKLLIKYASNLESG	61
Qy	516	VPDRFGSGYGTDPLTLISSVQAEDAAYVPCQDYNSPTTFGGGTGLETRADAAPTCSI	575
Db	62	VPARFSGSGSDTFLNIHPVEDEATYYCQHSEWLPIYTFGGGTGLETRADAAPTCSI	121
Qy	576	FPPSEQULTSGGASVGCFLNNYPKDINVKWKIDGSRQNGVLNSWTDDQSKDSTYSMS	635
Db	122	FPPSEQULTSGGASVGCFLNNYPKDINVKWKIDGSRQNGVLNSWTDDQSKDSTYSRSS	181
Qy	636	TLLTLTKDEYERHNSYTCEATHKTSTSPIVKSFRNE	671
Db	182	TLLTLTKDEYERHNSYTCEATHKTSTSPIVKSFRNE	217
RESULT 45			
US-09-170-769A-8			
; Sequence 8, Application US/09170769A			
; Patent No. 644206			
; GENERAL INFORMATION:			
; APPLICANT: THE SCRIPPS RESEARCH INSTITUTE			
; APPLICANT: LETURCO, Didier			
; APPLICANT: MORIATRY, Ann			
; APPLICANT: ULEVITCH, Richard			
; APPLICANT: TOBIAS, Peter			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING CD14 MEDIAT			
; FILE REFERENCE: SCRIPL1440-3			
; CURRENT APPLICATION NUMBER: US/09/170,769A			
; CURRENT FILING DATE: 1998-10-13			
; PRIOR APPLICATION NUMBER: US 08/070,160			
; PRIOR FILING DATE: 1993-05-28			
; NUMBER OF SEQ ID NOS: 26			
; SOFTWARE: PatentIn version 3.0			

```

: APPLICANT: TUBIAC, FELI
:
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITING CD14 MEDIATED CELL ACT
:
: FILE REFERENCE: SCRIPT1140-3
:
: CURRENT APPLICATION NUMBER: US/09/170,769A
:
: CURRENT FILING DATE: 1998-10-13
:
: PRIOR APPLICATION NUMBER: US 08/070,160
:
: PRIOR FILING DATE: 1993-05-28
:
: NUMBER OF SEQ ID NOS: 26
:
: SOFTWARE: PatentIn version 3.0

```

```

; SEQ ID NO 8
; LENGTH: 211
; TYPE: prt
; ORGANISM: Murine
US-09-170-769A-8

Query Match          25.5%; Score 898; DB 4; Length 211;
Best Local Similarity 81.4%; Pred. No. 5.3e-53;
Matches 171; Conservative 15; Mismatches 24; Indels 0; Gaps 0;

Qy 462 MTQPTSLIVSAGDRVITCKASQSVNDVAVYQKQSPKLLISYTSRYAGVPDRS 521
Db 1 MTQPTSSLSASLGRVITCKASQSDIKNYLWYQGGPGTVKLIYYTSLRHSGVPSRF 60
Qy 522 GSGYGTDTLTITSSVOAEDAAYVFCQDYNPSPTFGGKTLEIKRADAAAPTIVSIFPPS 581
Db 61 GSGGTDYSLTISNLEQDFATYFCQRGDTLPWTFGGGKTLEIKTADAAAPTIVSIFPPS 120
Qy 582 QLTSGGASVVCFLNNFYPKDINVKWKIDGSRQGVLSNWTQDQSKDSTYSMSSTLT 641
Db 121 QLTSGGASVVCFLNNFYPKDINVKWKIDGSRQGVLSNWTQDQSKDSTYSMSSTLT 180
Qy 642 DEYERHNSYTCETHKTSPIVKSFNRE 671
Db 181 DEYERHNSYTCETHKTSPIVKSFNRE 210

RESULT 46
PCT-US94-14106-55
; Sequence 55, Application PC/TUS9414106
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: Process for Generating Specific Antibodies
; NUMBER OF SEQUENCES: 61
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII (text)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/14106
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEQ ID NO: 55:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US94-14106-55

Query Match          25.5%; Score 897.5; DB 5; Length 223;
Best Local Similarity 79.4%; Pred. No. 6.1e-53;
Matches 177; Conservative 12; Mismatches 31; Indels 3; Gaps 2;

Qy 1 EVLQQSGPDLIVKPGASVKISKAGSYSTGYMHWKQSPGKGLWIGRINPNNGVTLY 60
Db 1 EVLQQSGAELMMPGASVKISKATGYTLSSWLEWYKQSPGHLWIGELIFCGSGSAHY 60
Qy 61 NQKFKDRLTVDKSSSTAYMELRSLTSEDSAVYYCARSTMITNY--YMDYWGQGTSTV 118
Db 61 NEKFKGKATFTVDTSSNTAYMQLSSLTSEDSAVYYCARGD--YGNVGDYFDYWGQGLTV 119
Qy 119 SSATKTPPSVYPLAGSAQAQNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFP 178
Db 120 FSAKTPPSVYPLAGSAQAQNSMVTGLCLVKGYLPEVTVTWNSGSLSSGVHTFP 179
Qy 179 SDLYTLSSSVTPSPSTPSETVTCNVAHPASSTKVDRKIVPRD 221
Db 180 SDLYTLSSSVTPSPSTPSETVTCNVAHPASSTKVDRKIVPRD 222

RESULT 47
```

```

US-08-737-129A-8
; Sequence 8, Application US/08737129A
; Patent No. 5885816
; GENERAL INFORMATION:
; APPLICANT: Ikuo FUJII et al.
; TITLE OF INVENTION: CATALYTIC ANTIBODIES ENANTIOSELECTIVELY
; TITLE OF INVENTION: HYDROLYSING AMINO ACID ESTER DERIVATIVES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack
; STREET: 805 Fifteenth Street, N.W., #700
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: Wordperfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/737,129A
; FILING DATE: No. 5885816ember 15, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Warren M. Cheek, Jr.
; REGISTRATION NUMBER: 33,367
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-371-8850
; TELEFAX:
; TELEX:
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 215 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-737-129A-8

Query Match          25.4%; Score 895; DB 2; Length 215;
Best Local Similarity 80.2%; Pred. No. 8.5e-53;
Matches 170; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

Qy 460 IVMTQPTSLIVSAGDRVITCKASQSVNDVAVYQKQSPKLLISYTSRYAGVPDR 519
Db 2 LVMTQPTSSMYASLGERVITCKASQDINIYLSWFQOKPGKSPKALIVRTNGLVDGVPSR 61
Qy 520 FSGSGYGTDTLTITSSVOAEDAAYVFCQDYNPSPTFGGKTLEIKRADAAAPTIVSIFPPS 579
Db 62 FSGSGSQDYSLTISSEYEDMGVIYCIQYDFEFTYFGGKTLEIKRADAAAPTIVSIFPPS 121
Qy 580 SEQLTSGASVVCFLNNFYPKDINVKWKIDGSRQGVLSNWTQDQSKDSTYSMSSTLT 639
Db 122 SEQLTSGASVVCFLNNFYPKDINVKWKIDGSRQGVLSNWTQDQSKDSTYSMSSTLT 181
Qy 640 TKDEYERHNSYTCETHKTSPIVKSFNRE 671
Db 182 TKDEYERHNSYTCETHKTSPIVKSFNRE 213

RESULT 48
US-08-442-542-18
; Sequence 18, Application US/08442542
; Patent No. 5685600
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine B.
; APPLICANT: Kozziel, Michael G.
; TITLE OF INVENTION: Antibodies which Bind to Insect Gut
```


;; TITLE OF INVENTION: Proteins and their Use
;; NUMBER OF SEQUENCES: 49
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Ciba-Geigy Corporation
;; STREET: 7 Skyline Drive
;; CITY: Hawthorne
;; STATE: NY
;; COUNTRY: USA
;; ZIP: 10532
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/442,542
;; FILING DATE: 16-MAY-1995
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/267,641
;; FILING DATE: 28-JUN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Spruill, W. Murray
;; REGISTRATION NUMBER: 32,943
;; REFERENCE/DOCKET NUMBER: CGC 1750
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 919-541-8615
;; TELEFAX: 919-541-8689
;; INFORMATION FOR SEQ ID NO: 18:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 599 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-442-542-18

Query Match 25.3%; Score 892.5; DB 1; Length 599;

Best Local Similarity 32.4%; Pred. No. 4.7e-52;

Matches 219; Conservative 73; Mismatches 118; Indels 265; Gaps 16;

QY 1 EVLOQSGPDLVKPGASVKISKASGYSTGYGMHWKQSPCKGLEWIGRI--NPNNGVT 58
DB 155 QVKLESQGGGLVQPKGSLKSCAASGFTFNFMVNRQAPGKGLWEVARIRSKSNVAT 214
QY 59 LYNQKFKDKATLTVDKSSSTAYMELRSITSEDSVAVYICARSTMITNYVMDYWGQTSVTV 118
DB 215 SYGDSVKDRFTVSRDSSQSMFYLQNNLKTEDTAMYYCVR---VYVGANDYWGQTSVTV 271
QY 119 SSAKTPPSVYPLAPG--SAAQTSNMVTLGCLVKGYFPPEPVTVWNSGLSGVHTFPVAV 176
DB 272 SSAKTPPSVYPLAPGSRSAQAQTSNMVTLGCLVKGYFPPEPVTVWNSGLSGVHTFPVAV 331
QY 177 LOSDLVTLSSSVTPSPSTVTCNVAHPASSTKVDKKIIVPRDSGGSPSEKSEINEKD 236
DB 332 LOSDLVTLSSSVTPSPSTVTCNVAHPASSTKVDKKIIVPRDGG----- 378
QY 237 LRKKSQLOQTALGNLQKIYYNSKAITSSEKSADQFLTNLTLLKFGFTGHPWYNLLVDL 296
DB 379 ----CKPCICTVPEVSSVFIFPPK-----PKDVLITL----- 407
QY 297 GSTAATSEYEGSSVDLYCAYGYCAGGTPNKTACMYGGVTUHDNNRLTEKKVPINLWI 356
DB 408 -----TP-KVTCVVVDIS-----KDDPEVQFSWFV 431
QY 357 DGKQTTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-VNSDSFGCKVGORG 407
DB 432 D----DVEVHTAQTPREQFNSTRSRVSELPIMHODLWNGREFKCRVNSAFAFPPIEK- 486
QY 408 LIVFHSSEGSTVSYDLFDAQQGYPTDLLRIYRDNTITTSLSLSLSLYLYTTSIVMTQPT 467
DB 487 -----TISK----- 491
QY 468 SLLVSNAGDRVTTCKASQSVSNDVAVYQOKPGQSPKLLISYTSRYAGVPDRFSGSGYGT 527

DB 492 ----- 491
QY 528 DFTLTLSVQAEDAAYFCQDYNSPPTFGGTGKLEIKRADAPTYSIFPPSSEOLTSGG 587
DB 492 -----KGRKAPQVYTIPTPKQMAKDK 514
QY 588 ASVVCFLANNFYPKDINVKWKIDGSRQN-----GVLNSWTDQDSKDSYVSMSSLTLTAKD 642
DB 515 VSLTCMITDFFPEDITVEQWNGQPAENYKNTQPIMNT-----NGSYFVYKLVNQKS 567
QY 643 EYERHNSYTCETHK 657
DB 568 NWEAGNTFTCSVLHE 582
RESULT 49
US-08-765-469-18
; Sequence 18, Application US/08765469
; Patent No. 6069301
; GENERAL INFORMATION:
; APPLICANT: Carozzi, Nadine B.
; APPLICANT: Koziel, Michael G.
; TITLE OF INVENTION: Antibodies which Bind to Insect Gut
; TITLE OF INVENTION: Proteins and their Use
; NUMBER OF SEQUENCES: 49
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ciba-Geigy Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: NY
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/765,469
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/267,641
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Spruill, W. Murray
; REGISTRATION NUMBER: 32,943
; REFERENCE/DOCKET NUMBER: CGC 1750
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8615
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 599 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-765-469-18

Query Match 25.3%; Score 892.5; DB 3; Length 599;

Best Local Similarity 32.4%; Pred. No. 4.7e-52;

Matches 219; Conservative 73; Mismatches 118; Indels 265; Gaps 16;

QY 1 EVLOQSGPDLVKPGASVKISKASGYSTGYGMHWKQSPCKGLEWIGRI--NPNNGVT 58
DB 155 QVKLESQGGGLVQPKGSLKSCAASGFTFNFMVNRQAPGKGLWEVARIRSKSNVAT 214
QY 59 LYNQKFKDKATLTVDKSSSTAYMELRSITSEDSVAVYICARSTMITNYVMDYWGQTSVTV 118
DB 215 SYGDSVKDRFTVSRDSSQSMFYLQNNLKTEDTAMYYCVR---VYVGANDYWGQTSVTV 271
QY 119 SSAKTPPSVYPLAPG--SAAQTSNMVTLGCLVKGYFPPEPVTVWNSGLSGVHTFPVAV 176

Db 272 SSAKTPPSVYPLAPGSRSAQAQNSMVTGLCLVKGYPPEPTVTWNSGLSSGVHTFPAY 331
Qy 177 LOSDLTYLSSSVYVPSSTWPSSTVTCNVAHPASSTKVDKIVPRDSGCPSEKSEINEKD 236
Db 332 LOSDLTYLSSSVYVPSSTWPSSTVTCNVAHPASSTKVDKIVPRDCG----- 378
Qy 237 LRKKSELOAGTALGNLQIYYNSKAITSSSEKSAQDFTLTLFKGFTTGHWPYNDLLVDL 296
Db 379 ---CKPCICIVPEVSSVFIPEPK-----PKDVLITL----- 407
Qy 297 GSTAATSEYEGSSVDLYGAYGYOCAGTGNKTAQMYGGVTLHDNNRLTEKKYVINLWI 356
Db 408 -----TP-KVTCVVVDIS-----KDDPEVQFSWFV 431
Qy 357 DGKQTTVPIDKVKTSKKEV-----TVQELDLQARHVLHCK-FGL-YNSDSFGKVGORG 407
Db 432 D-----DVEHTAQTQPREEQNFSTRSELPIHQDLNGKEFKCRVNSAFAPIEK- 486
Qy 408 LIVFHSSEGSTVSVDFDAQGYPTLLRIYRDNTTISSTSLISLYLYTTTSIVMTQTPT 467
Db 487 -----TISKT----- 491
Qy 468 SLLVSAGDRVTTICKASQSVNDVAVYQKPGQSPKLLISYTSRYAGVPRFSGSGYGT 527
Db 492 ----- 491
Qy 528 DFTLTSSVQAEADAAYFCQODYNSPPTFGGKLEIKRADAAAPTSTVIFPPSSQLTSGG 587
Db 492 -----KGRPKAPQVYTIPTPPKEQMAKDK 514
Qy 588 ASVVCFLNNFYKPDINVKWKIDGSRQV-----GVLSNWTQDQSKDSTYSMSSTLTITKD 642
Db 515 VSLTCMTIDFPEDITVEMQNGQPAENYKNTQPMNT-----NGSYFYVSKLNVQKS 567
Qy 643 EYERHNSYTCEATHK 657
Db 568 NWEAGNTFTCSVLHE 582

RESULT 50
US-09-423-439-58
; Sequence 58, Application US/09423439
; Patent No. 6339070
; GENERAL INFORMATION:
; APPLICANT: EMERY, Stephen Charles
; TITLE OF INVENTION: CHEMICAL COMPOUNDS
; NUMBER OF SEQUENCES: 60
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pillsbury Winthrop, L.L.P.
; STREET: 1100 New York Ave., N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/423.439
; FILING DATE: 09-No. 6339070-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB98/01294
; FILING DATE: 05-MAY-1998
; APPLICATION NUMBER: GB 9709421.3
; FILING DATE: 10-MAY-1997
; INFORMATION FOR SEQ ID NO: 58:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 235 amino acids

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 58:
US-09-423-439-58

Query Match 25.3%; Score 889.5; DB 4; Length 235;
Best Local Similarity 77.1%; Pred. No. 2.2e-52;
Matches 172; Conservative 19; Mismatches 31; Indels 1; Gaps 1;

Qy 449 LSISLYLYTTSIVMTQTPTSLLVSAGDRVTTICKASQSVNDVAVYQKPGQSPKLLISY 508
Db 13 ISASVIMSRGQTVLSQSPAILSASPGKVTMTCRASSVTY-IHWYQKPGGSPKSWIYA 71
Qy 509 TSSRYAGVPRDFSGGYCTDFTLTSSVQAEADAAYFCQODYNSPPTFGGKLEIKRAD 568
Db 72 TSNLASGVPARFSGSGGTSTSLTISRVEADAATYTCQHWSSSKPPTFGGKLEIKRAD 131
Qy 569 AAPTSTVIFPPSSQLTSGGASVVCFLNNFYKPDINVKWKIDGSRQVGLNSWTQDQSKD 628
Db 132 AAPTSTVIFPPSSQLTSGGASVVCFLNNFYKPDINVKWKIDGSRQVGLNSWTQDQSKD 191
Qy 629 STYSMSSTLTITKDEYERHNSYTCEATHKSTSPIVKSFNRNE 671
Db 192 STYSMSSTLTITKDEYERHNSYTCEATHKSTSPIVKSFNRNE 234

Search completed: January 8, 2003, 11:59:34
Job time : 30.0486 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:58:32 ; Search time 12.6232 Seconds
(without alignments)
1032.821 Million cell updates/sec

Title: US-09-900-766-1

Perfect score: 3522

Sequence: 1 EVQLQQSGPDLVKPGASVKI.....EATHKTSPIVKSFNRES 672

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/FCI_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/FCIUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1257	35.7	569	US-09-807-721-2	Sequence 2, Appli
2	1072	30.4	248	US-09-870-759-16	Sequence 16, Appli
3	1028	29.2	597	US-09-903-327A-11	Sequence 11, Appli
4	1024.5	29.1	613	US-09-903-327A-14	Sequence 14, Appli
5	1023.5	29.1	456	US-09-903-327A-2	Sequence 2, Appli
6	1023.5	29.1	493	US-09-903-327A-13	Sequence 13, Appli
7	1023.5	29.1	510	US-09-903-327A-12	Sequence 12, Appli
8	1022.5	29.0	438	US-09-903-327A-6	Sequence 6, Appli
9	935	26.5	257	US-09-870-759-8	Sequence 8, Appli
10	912	25.9	238	US-09-903-327A-4	Sequence 4, Appli
11	855.5	24.3	473	US-09-910-059-131	Sequence 131, App
12	807	22.9	448	US-09-917-410-6	Sequence 6, Appli
13	782	22.5	461	US-09-249-011A-24	Sequence 24, Appli
14	789	22.4	473	US-09-828-995B-20	Sequence 20, Appli
15	789	22.4	729	US-09-825-012-52	Sequence 52, Appli
16	789	22.4	730	US-09-825-012-49	Sequence 49, Appli
17	789	22.4	731	US-09-825-012-46	Sequence 46, Appli
18	789	22.4	739	US-09-825-012-61	Sequence 61, Appli
19	789	22.4	740	US-09-825-012-58	Sequence 58, Appli

20	789	22.4	741	10	US-09-825-012-55	Sequence 55, Appli
21	785.5	22.3	470	10	US-09-859-053-28	Sequence 28, Appli
22	780	22.1	230	10	US-09-995-693-2	Sequence 2, Appli
23	779.5	22.1	476	10	US-09-747-669-3	Sequence 3, Appli
24	776.5	22.0	232	10	US-09-754-998-2	Sequence 2, Appli
25	776.5	22.0	468	9	US-09-828-995B-5	Sequence 5, Appli
26	759.5	21.6	470	9	US-09-828-995B-11	Sequence 11, Appli
27	756	21.5	450	10	US-09-811-384-12	Sequence 12, Appli
28	740	21.0	240	10	US-09-799-514-8	Sequence 8, Appli
29	728	20.7	237	10	US-09-940-166A-6	Sequence 6, Appli
30	726	20.6	214	10	US-09-940-166A-2	Sequence 2, Appli
31	726	20.6	214	10	US-09-811-384-11	Sequence 11, Appli
32	726	20.6	220	10	US-09-995-693-1	Sequence 1, Appli
33	725.5	20.6	300	10	US-09-940-166A-7	Sequence 7, Appli
34	722.5	20.5	515	10	US-09-825-012-66	Sequence 66, Appli
35	722.5	20.5	517	10	US-09-825-012-38	Sequence 38, Appli
36	722.5	20.5	519	10	US-09-825-012-76	Sequence 76, Appli
37	722.5	20.5	519	10	US-09-825-012-80	Sequence 80, Appli
38	722.5	20.5	521	10	US-09-825-012-71	Sequence 71, Appli
39	722.5	20.5	525	10	US-09-825-012-85	Sequence 85, Appli
40	722.5	20.5	527	10	US-09-825-012-43	Sequence 43, Appli
41	722.5	20.5	529	10	US-09-825-012-95	Sequence 95, Appli
42	722.5	20.5	531	10	US-09-825-012-90	Sequence 90, Appli
43	718.5	20.4	232	10	US-09-811-384-10	Sequence 10, Appli
44	718.5	20.4	241	10	US-09-940-166A-1	Sequence 1, Appli
45	718.5	20.4	449	10	US-09-736-371B-21	Sequence 21, Appli

ALIGNMENTS

RESULT 1
US-09-807-721-2
; Sequence 2, Application US/09807721
; Patent No. US20020174453A1
; GENERAL INFORMATION:
; APPLICANT: AUBURN UNIVERSITY
; APPLICANT: UNIVERSITY OF CENTRAL FLORIDA
; TITLE OF INVENTION: PRODUCTION OF ANTIBODIES IN TRANSGENIC PLASTIDS
; FILE REFERENCE: 1463-PCT-US-00
; CURRENT APPLICATION NUMBER: US/09/807,721
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US01/06274
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-807-721-2

Query Match	35.7%	Score 1257;	DB 9;	Length 669;
Best Local Similarity	41.6%;	Pred. No. 3.6e-47;		
Matches 308;	Conservative 87;	Mismatches 201;	Indels 144;	Gaps 23
QY	2	VLQSQSGPDLVKPGASVKISKASGYFTGYHWHVQKSPGKLEWIGRINPNNGVTLYN	61	
Db	3	VLQSQSGPDLVKPGASVKISKASGYFTDYNIHWVQKSRGKSLWIGYIYPNGNTYYN	62	
QY	62	QKFKDKATLTVDKSSITAYMELRSLTSEDSAVVYCARSTMTNVMYMGQGTSTVTSSA	121	
Db	63	QKFKNKATLTVDNSSTSAYMELRSLTSEDSAVVYCAT-----YFDYMGQGTTLTVSSA	115	
QY	122	KTTTPSVYPLAPGSAOATNSMWTLGCLVKGYP-EPVTVWNSGSLSSGVHFFPAV--LQ	178	
Db	116	SPTSPKVFPLSLDSTPDQGNVV-VACLVGQFPQEPFLSVTWSGQNVARTNPPSQDAS	174	
QY	179	SDLTYTLSSSVTPSPSWP-SETVTCNVAHPASSTKVDDKIVPRDSGGGSEKSEINEKDL	237	
Db	175	GDIYTTSSQLTLPATCPDGKSVTCVKHYHTNSSQ-DVTVPGRVPPPP-----PCC	224	
QY	238	RKKSELOQTALGNLKOIYYYNKAITTSEKSAQDLTNTLLFKGFTTGHVWYNDLLVDLG	297	


```

; TYPE: PROT
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (0)---(0)
; OTHER INFORMATION: DAV-1 heavy chain, penton base monoclonal antibody
; US-09-903-327A-2

Query Match          29.1%; Score 1023.5; DB 9; Length 456;
Best Local Similarity 37.4%; Pred. No. 2.2e-37;
Matches 249; Conservative 51; Mismatches 111; Indels 255; Gaps

Qy 1 EVQLQSGDPLVKPGASVKISKASGYSFTQYIMHWKQSPGKLEWIGRINPNNGVTLY 60
    |||||
Db 20 EVQLQSGPELVKPGASVKISKASGYFTDYNNHWVKQSHGKSLIEWIGYIPYRGCTGY 79
    |||||

Qy 61 NQPKFKATLTVDKSGTAYMELRSLTSEDSAVYICARSTMITNYMVDYWGQCTSVTVSS 12
    |||||
Db 80 NQPKFKATLTITDSSNTAYMELRSLTSDAGAVYICARG-----IAYWGQGLTIVTVA 13
    |||||

Qy 121 AKTTPPSVYPLAPGSAQTNSMVTLCGLVKGYFPPEYVTVWNSGSLSSGVHFPFPAVLQSD 18
    |||||
Db 133 AKTTPSVYPLAPGSAQTNSMVTLCGLVKGYFPPEYVTVWNSGSLSSGVHFPFPAVLQSD 19
    |||||

Qy 181 LYTLSSTVVPSPSTWPSPTVTCNVAHPASSTKVDKKIIVPRDSGGSPSEKSEETNEKDLRKK 24
    |||||
Db 193 LYTLSSTVVPSPSTWPSPTVTCNVAHPASSTKVDKKIIVPRDCG----- 23
    |||||

Qy 241 SELQGTALGNLKQIYYNSKAITSSSEKSAQDQFLNTLLFKGFTGHPWYNDLLVDLGGSTA 30
    |||||
Db 236 CKPCICTVPEVSSVFIFFPK-----PKDVLITL----- 26
    |||||

Qy 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKACMYGGVTLHDNNRLTBEKKVPINLMIDGKQ 36
    |||||
Db 265 -----TP-KVTCCVVVDIS-----KDDPEVQVSWFVD--- 28
    |||||

Qy 361 TTVPIDKVTSKREV-----TVQELDLQARHYLHGK-EGL-YNSDSFGGKVQGRGLIVF 41
    |||||

```

Db 290 -DVEVHTAQTPREEQFNSTFRSVSELPIMHQDWLNGEFKCRVNSAAFPAPIEK----- 343
Qy 412 HSSEGSTVSYDLFDAQGQYPTDLLRIYRDNNTTISSTLSISLYLTTISVMTQTPTSLLV 471
Db 344 -----TISK-----
Qy 472 SAGDRVTITCKASQSVNDVAVYQKPGQSKLLISYSSRYAGVDPDRFSGSGYGTDFTL 531
Db 349 -----
Qy 532 TISSVOAEDAAVYFCQDYNSPPTFGGKLEIKRADAAPTIVSIPPSSEQLTSGGASV 591
Db 349 -----KGRKPAQVYTIPTPPKEQMAKDKVSLT 375
Qy 592 CFLNFPYKIDINVKKIDGSRQNGVLSWTDQDSDKSTYSMSSTLTTLTKDEYERHNSYT 651
Db 376 CMITDFPEDITVEQWNGQPAEN-YKNTQPIMDT-DGSYFYVSKLVNOKSNWEAGNTFI 433
Qy 652 CEATHK 657
Db 434 CSVLHE 439
RESULT 6
US-09-903-327A-13
; Sequence 13, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erquang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 493
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and EGF mature peptide
US-09-903-327A-13

Query Match 29.1%; Score 1023.5; DB 9; Length 493;
Best Local Similarity 37.4%; Pred. No. 2.3e-37;
Matches 249; Conservative 51; Mismatches 111; Indels 255; Gaps 14;
Qy 1 EVOLQSQGDLVKPGASVKISKASGYSTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db 20 EVOLQSQGPELVKPGASVKISKASGYSTGYIMHWKQSHGKSLWIGIYIPYKGGTY 79
Qy 61 NQKFKDKATLTVDKSSYAYMELSLTSDSAVYICARSTMTNTVMYWGOGTSVTYSS 120
Db 80 NQKFKSKATLTDDSSNTAYMELSLTSDSAVYICARG-----IAYWGGTGLTVTVA 132
Qy 121 AKTTPPSVYPLAPGSAATNSMTLGLCLVKGYPPEPVTWNSGSLSSGVHTFPVAVLQSD 180
Db 133 AKTTPPSVYPLAPGSAATNSMTLGLCLVKGYPPEPVTWNSGSLSSGVHTFPVAVLQSD 192
Qy 181 LYTSSSVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDSGGPSEKSEINEKDLRKK 240
Db 193 LYTSSSVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDCG-----
Qy 241 SELOGTALGNLQIYYNYSKAITSSSEKADQFLTWLLFGFTGHPWYNDLLDLGSTA 300
Db 236 CRPCICTVPEVSVFIPFPK-----PKDVLITL-----
,

Qy 301 ATSEYEGSVLDLYGAYGYQCAGGTPNKACMYGGVTLHDNNRLTEKKVPIINLWIDGQ 360
Db 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD--- 289
Qy 361 TTVPIDKYTKSKY-----TVQELDLQARHYLHGK-FGL-YNSDSFGKQVQRLIVF 411
Db 290 -DVEVHTAQTPREEQFNSTFRSVSELPIMHQDWLNGEFKCRVNSAAFPAPIEK----- 343
Qy 412 HSSEGSTVSYDLFDAQGQYPTDLLRIYRDNNTTISSTLSISLYLTTISVMTQTPTSLLV 471
Db 344 -----TISK-----
Qy 472 SAGDRVTITCKASQSVNDVAVYQKPGQSKLLISYSSRYAGVDPDRFSGSGYGTDFTL 531
Db 349 -----
Qy 532 TISSVOAEDAAVYFCQDYNSPPTFGGKLEIKRADAAPTIVSIPPSSEQLTSGGASV 591
Db 349 -----KGRKPAQVYTIPTPPKEQMAKDKVSLT 375
Qy 592 CFLNFPYKIDINVKKIDGSRQNGVLSWTDQDSDKSTYSMSSTLTTLTKDEYERHNSYT 651
Db 376 CMITDFPEDITVEQWNGQPAEN-YKNTQPIMDT-DGSYFYVSKLVNOKSNWEAGNTFI 433
Qy 652 CEATHK 657
Db 434 CSVLHE 439
RESULT 7
US-09-903-327A-12
; Sequence 12, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erquang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TAR
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Fusion protein with N-terminal portion of DAV-1 heavy chain
; OTHER INFORMATION: and IGF-1 mature peptide
US-09-903-327A-12
Query Match 29.1%; Score 1023.5; DB 9; Length 510;
Best Local Similarity 37.4%; Pred. No. 2.4e-37;
Matches 249; Conservative 51; Mismatches 111; Indels 255; Gaps 14;
Qy 1 EVOLQSQGDLVKPGASVKISKASGYSTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db 20 EVOLQSQGPELVKPGASVKISKASGYSTGYIMHWKQSHGKSLWIGIYIPYKGGTY 79
Qy 61 NQKFKDKATLTVDKSSYAYMELSLTSDSAVYICARSTMTNTVMYWGOGTSVTYSS 120
Db 80 NQKFKSKATLTDDSSNTAYMELSLTSDSAVYICARG-----IAYWGGTGLTVTVA 132
Qy 121 AKTTPPSVYPLAPGSAATNSMTLGLCLVKGYPPEPVTWNSGSLSSGVHTFPVAVLQSD 180
Db 133 AKTTPPSVYPLAPGSAATNSMTLGLCLVKGYPPEPVTWNSGSLSSGVHTFPVAVLQSD 192
Qy 181 LYTSSSVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDSGGPSEKSEINEKDLRKK 240
Db 193 LYTSSSVTPSPSTWPSSETVTCNVAHPASSTKVKDKKIVPRDCG-----PKDVLITL-----
,

Db 193 LYTSSVTPSPSTWSETVTCNVAHPASSTKVKKIVPRDCG----- 235
QY 241 SELQGTALGNLKOIYYNSKAITSEKSDAQFLNTLTFKGFETHGPHWYNDLLVDLGSTA 300
Db 236 CKPCICTVPEVSSVFPPK-----PKDVLITL----- 264
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEEKVVPINLWIDGKQ 360
Db 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD--- 289
QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVORGLIYF 411
Db 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHQDWLNGKEKCRVNSAAPAPIEK----- 343
QY 412 HSSEGSTVSYDLFDAOGQYPTDLLRIYRDNNTTISSTLSLSISLYLYTTSIVMTQPTPTSLLV 471
Db 344 -----TISKT----- 348
QY 472 SAGDRVTITCKASQSVNSDVAMVYQKPGOSP KLLISYTSRRYAGVDPDRFSGSGYGTDFTL 531
Db 349 ----- 348
QY 532 TISSVQAEADAAVYFCQDYNPPTFGGTTKLEIKRADAAPTYSIFPPSSEQLTSGCASV 591
Db 349 -----KGRPKAPQVYTIPTPPKEQMAKDKVSLT 375
QY 592 CFLNFPKIDINVKIKIDGSRONGVLSWTDQSDKSTYSMSSTLTLTQKDEYERHNSYT 651
Db 376 CMITDFEPEDITVEQWNGQPAEN-YKNTQPIMDT-DGSYFYVYKLNQVKSWEAGNTFI 433
QY 652 CEATHK 657
Db 434 CSVLHE 439

RESULT 8
US-09-903-327A-6
; Sequence 6, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erguang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22908-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 438
; TYPE: PRT
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (0)...(0)
; OTHER INFORMATION: Portion of DAV-1 heavy chain used for fusion protein
; OTHER INFORMATION: bifunctional antibody
US-09-903-327A-6

Query Match 29.0%; Score 1022.5; DB 9; Length 438;
Best Local Similarity 37.4%; Pred. No. 2.3e-37;
Matches 249; Conservative 50; Mismatches 111; Indels 255; Gaps 14;
QY 1 EVLOQSGDPLVKPGASVKISKASGYSTGYIMHWKVGSPKGLIEWIRNPNNGVTLY 60
Db 20 EVLOQSGPELVKPGASVKISKASGYTFTDYNMHWKVGSHGKSLIEWIGYIYKGGTCY 79
QY 61 NQFKDKATLTVDKSTTAYMELRSLTSDSAVYCYCARSTMTITNYWVGOGTSTVSS 120

Db 80 NQFKSKATLTDTSSSNTAYMELRSLTSDSAVYCYCARG-----IAYWGGTTLTVSA 132
QY 121 AKTTTPSVYPLAPGSAQTNSMVTLCGLVKGYFPPEVTVTNWNSGSLSSGVHTFPFPAVLQSD 180
Db 133 AKTTTPSVYPLAPGSAQTNSMVTLCGLVKGYFPPEVTVTNWNSGSLSSGVHTFPFPAVLQSD 192
QY 181 LYTSSVTPSPSTWSETVTCNVAHPASSTKVKKIVPRDGGPSPSEKSEINEKDLRKK 240
Db 193 LYTSSVTPSPSTWSETVTCNVAHPASSTKVKKIVPRDCG----- 235
QY 241 SELQGTALGNLKOIYYNSKAITSEKSDAQFLNTLTFKGFETHGPHWYNDLLVDLGSTA 300
Db 236 CKPCICTVPEVSSVFPPK-----PKDVLITL----- 264
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEEKVVPINLWIDGKQ 360
Db 265 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD--- 289
QY 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVORGLIYF 411
Db 290 -DVEVHTAQTPREQFNSTFRSVSELPIMHQDWLNGKEKCRVNSAAPAPIEK----- 343
QY 412 HSSEGSTVSYDLFDAOGQYPTDLLRIYRDNNTTISSTLSLSISLYLYTTSIVMTQPTPTSLLV 471
Db 344 -----TISKT----- 348
QY 472 SAGDRVTITCKASQSVNSDVAMVYQKPGOSP KLLISYTSRRYAGVDPDRFSGSGYGTDFTL 531
Db 349 ----- 348
QY 532 TISSVQAEADAAVYFCQDYNPPTFGGTTKLEIKRADAAPTYSIFPPSSEQLTSGCASV 591
Db 349 -----KGRPKAPQVYTIPTPPKEQMAKDKVSLT 375
QY 592 CFLNFPKIDINVKIKIDGSRONGVLSWTDQSDKSTYSMSSTLTLTQKDEYERHNSYT 651
Db 376 CMITDFEPEDITVEQWNGQPAEN-YKNTQPIMDT-DGSYFYVYKLNQVKSWEAGNTFI 433
QY 652 CEATHK 656
Db 434 CSVLH 438

RESULT 9
US-09-870-759-8
; Sequence 8, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 257
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-870-759-8

Query Match 26.5%; Score 935; DB 9; Length 257;
Best Local Similarity 75.5%; Pred. No. 7.7e-34;
Matches 176; Conservative 21; Mismatches 36; Indels 0; Gaps 0;
QY 226 SEKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSEKSDAQFLNTLTFKGFETHG 285
Db 25 SEKSEINEKDLRKKSELOGTALGNLKOIYYNEKAKTENKESHQOFLQHTILFKGFFTN 84
QY 286 HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345

Db 85 HSWYNDLLVDFSDKIDVRYKGGKVDLYGAYYGYQACGTPNKTACMYGGVTLHDNRILT 144
Qy 346 EKKKPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHVLHGKFGLYNSDSFGGKVQ 405
Db 145 EKKKPINLWLDGKONTVPLETVKTKNKNVTVOELDLQARRYLQKYNLYNSDVFDDGVQ 204
Qy 406 RGLIVFHSSEGSVSYDLFADGQYPTDLLRIYRDNNTISSTLSLSISLYLT 458
Db 205 RGLIVFHTSTEPSVNYDLFGAQQNSNTLLRIYRDNKNTINSENMHIDIYLYTS 257

RESULT 10
US-09-903-327A-4
; Sequence 4, Application US/09903327A
; Patent No. US20020164333A1
; GENERAL INFORMATION:
; APPLICANT: Nemerow, Glen R.
; APPLICANT: Li, Erquang
; TITLE OF INVENTION: BIFUNCTIONAL MOLECULES AND VECTORS COMPLEXED THEREWITH FOR TARGET
; TITLE OF INVENTION: GENE
; TITLE OF INVENTION: DELIVERY
; FILE REFERENCE: 22508-1228
; CURRENT APPLICATION NUMBER: US/09/903,327A
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 09/613,017
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Mouse
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (0)...(0)
; OTHER INFORMATION: DAV-1 light chain, penton base monoclonal antibody
US-09-903-327A-4

Query Match 25.9%; Score 912; DB 9; Length 238;
Best Local Similarity 81.3%; Pred. No. 6.8e-33;
Matches 178; Conservative 9; Mismatches 28; Indels 4; Gaps 1;

Qy 457 TTSIVMTQPTTSLVYSGADRVTTITCKASQVSDND----VAVYQQRPGQSPKLLISYTSR 512
Db 19 TGDIVLTQSPASLVSLGQRATISCKASQSDYDGDYNNWYQQRPGPKLLIYAASNL 78
Qy 513 YAGVDRSGSGYGTDFLTITSSVAQDAADYFCQDYNPPTFGGKTKLEIKRADAAPT 572
Db 79 ESGIPARFSGSGGTDFLTNIHPVEEDAATYICQQTREDPWTFGGKTKLEIKRADAAPT 138
Qy 573 VSIFPPSSQLTSGGASVVCFLNNFYPKDINVKWKIDGSGRQGVLSNWTDDSDKSDSYS 632
Db 139 VSIFPPSSQLTSGGASVVCFLNNFYPKDINVKWKIDGSGRQGVLSNWTDDSDKSDSYS 198
Qy 633 MSSTLTLTKDEYRHNSYTCEATHKSTSTSPIVKSFNRR 671
Db 199 MSSTLTLTKDEYRHNSYTCEATHKSTSTSPIVKSFNRRNE 237

RESULT 11
US-09-910-059-131
; Sequence 131, Application US/09910059
; Patent No. US20020142359A1
; GENERAL INFORMATION:
; APPLICANT: Copley, Clive G
; APPLICANT: Edge, Michael Derek
; APPLICANT: Emery, Stephen Charles
; TITLE OF INVENTION: Monoclonal Antibody to CEA, Conjugates Comprising Said Antibody,
; TITLE OF INVENTION: Their Therapeutic use in an Adept System
; FILE REFERENCE: 1991-209
; CURRENT APPLICATION NUMBER: US/09/910,059
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 09/171,945

; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: PCT/GB97/01165
; PRIOR FILING DATE: 1997-04-29
; PRIOR APPLICATION NUMBER: GB 9703103.3
; PRIOR FILING DATE: 1997-02-14
; PRIOR APPLICATION NUMBER: GB9609405.7
; PRIOR FILING DATE: 1996-05-04
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 131
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: full-length human B7.1-murine ASB7 Fd fusion
US-09-910-059-131

Query Match 24.3%; Score 855.5; DB 10; Length 473;
Best Local Similarity 72.9%; Pred. No. 3.1e-30;
Matches 164; Conservative 24; Mismatches 34; Indels 3; Gaps 3;

Qy 1 EVOLQSGPDLVKPCASVKISKASGYSTGYMHWVKQSPCKGLEWIGRI-NPNNGVTL 59
Db 243 EYKLVESGGGLVQPGGSLRLSCATSGFTTDTYIMNMVRQPPGKALEWLGFIGNKANGYTT 302
Qy 60 -YNOKFKDKATLTVDKSTTAYMELRSLTSDSVAVYVCARSTMITNMYMDYWGQGTSTVTV 118
Db 303 EYSASVKGRFTISRDKSOSILYLQNTLRAEDSAYYCYTRDGLRFY-FDWGQGTTLTV 361
Qy 119 SSAKTPPSVYPLAPGSAATNSMVTGLGKLVKGYPEPPTVTWNSGSLSSGVHTFPVLQ 178
Db 362 SSAKTPPSVYPLAPGSAATNSMVTGLGKLVKGYPEPPTVTWNSGSLSSGVHTFPVLQ 421
Qy 179 SDLYTLSSSVTPSPSTWPTSETVTCNVAHPASSTKVDDKIVPRDSG 223
Db 422 SDLYTLSSSVTPSPSTWPTSETVTCNVAHPASSTKVDDKIVPRDCG 466

RESULT 12
US-09-917-410-6
; Sequence 6, Application US/09917410
; Patent No. US20020098183A1
; GENERAL INFORMATION:
; APPLICANT: MARTIN, Ulrich; HASELBECK, Anton; SCHUMACHER, Guenther;
; CO. Man S.
; TITLE OF INVENTION: ANTI-L-SELECTIN ANTIBODIES FOR PREVENTION OF
; MULTIPLE ORGAN FAILURE AFTER POLYTRAUMA AND FOR
; PREVENTION OF ACUTE ORGAN DAMAGE AFTER
; EXTRACORPOREAL BLOOD CIRCULATION
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Computer Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII, WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/917,410
; FILING DATE: 26-Jul-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/578,953
; FILING DATE: <Unknown>
; APPLICATION NUMBER: EP 95 114 969.9
; FILING DATE: 19-Sep-95
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. US20020098183A1man D.


```

Qy 480 TCRASQSVSNDVAWYQQKPGQSPKLLISYTSSRYAGVDPDRFGSGYGTDFTLTITSSVQAE 539
    || |         | : | | | : |
Db 340 KCKVS-----NKALPAP---IEKTISKAKGP-----363
Qy 540 DAAVYFCQDYNPPPTGGGTKLEIKRADAAPTVISFPSPSEQLTSGGASVVCFLNPF 599
    | | | | | | | | | | | | | | | | | | | | |
Db 364 -----REPOVYTLPPSRDELTKNQVSLTCLVKGFYP 394
Qy 600 KDINVKWKIDGSERQN-----GYLNSWTDDSKDSSTYSMSSTLTCLKDEVERHNSVTCEA 654
    || | |:| :|| | | | | | | | | | | | | | |
Db 395 SDIAVESMGQPENNYKTPPVLDLS-----DGSFELYSKLTVDKSRMQQGNVFCSV 447
Qy 655 THK 657
    | :
Db 448 MHE 450

RESULT 18
US-09-825-012-61
; Sequence 61, Application US/09825012
; Patent No. US2002012798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFc1 heavy chain - DNase I fusion
US-09-825-012-61

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 heavy chain - DNase I fusion
us-09-825-012-58

Query Match      22.4%; Score 789; DB 10; Length 740;
Best Local Similarity 31.5%; Pred. No. 3e-27;
Matches 209; Conservative 65; Mismatches 151; Indels 238; Gaps 17

Qy 1 EVQLQQSGPDLVKPGASVKISKASGYFTGYMHMWKQSPGKGLEWIGRINPNNGVTLY 60
   ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20 QVQLVQSGAEVKKPGASVKVSCKASGYFTSAYIEWVRQAPGKGLEWVGEILPGSNRSY 79

Qy 61 NQKRFDKATLVDSKSTTAYMELSLTSEDSNYYVCARSTMITNYMVDYWGOGTSVTYSS 120
   ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 80 NEKTKGRVTYTRDSTNTAYMELSLRSEDATVYCARSDYFAWEA - YWGOGTLVTVSS 137

Qy 121 AKTTPPSVYPLAPGSAAGTNSMTVLCGLVKGYFPPEPVTWNSGSLSSGVHFPFPAVLQSD 180
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 ASTKGPSVFPLAPSKSTSGGTAAALGCLVKDYFPPEPVTWNSGALTSQVHFPFPAVLQSS 197

Qy 181 -LYTLSSSVTPSPSTWSPSTVTCNVAHPASSTKYDKKIVPRDSGGPSEKSEINEKDLRK 239
   ||| ||| ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 GLYSLSSVVTVPSSSLGTQTYICNVNHRPSNTKVDKVEPKSC - - - - - 240

Qy 240 KSELQGTALGNLKQIIYYNSKAITSSEKSAQDLFTNTLLKFGFTGHPWYNDLLVDLGST 299
   ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 - - - - - DKHTCCPPAPPELLCGPSVF - LFPKP - KDTLM - - - - - 272

Qy 300 AATSEYEGSSVDLYGAYGYOCAGTTPNKTACMGYGVTLHDNNRLTEKKVPINLWIDGK 359
   ||| ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 273 - - - - - ISRTPEVT - CVVVDDVS - - - - - HEDPEVAFNNVVDG - 301

```

Qy	360	QTTVPIDKVTSKSKEVTVQELDLQARRHYLHGKFGELGNSDFGKGKVGORGLIVFHSSEGSTV	4119
Db	302	---VEVHNAKTKPREQ-----YNSTY---RVSVLTVLHLQDWLNGK	337
Qy	420	SYDLFDAQGOYPDTLLRIYRDNTTISTSLISLYLYTTSIVMTQPTSLLSVAGDRVTI	479
Db	338	EY-----	339
Qy	480	TKKASQSVSNDVAMYOQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTLTLTSSVQAE	539
Db	340	KCKVS-----NKALPAP---LEKTSKAKGP-----	363
Qy	540	DAAVYFCQDYNPPPTFGGCTKLEIKRADAPVTSTFPSPSSSEOLTSGGASVVCFLNNFYP	599
Db	364	-----REPOVYTLPPSRDELTKNQVSLTCLVKGFYP	394
Qy	600	KDINVKWKIDGSERON-----GYLNSWTQDSKDSYTSMSSTLTLTPKDEYERHNSVTCEA	654
Db	395	SDIAVENESGQPNENYKTTTPVLDS-----DGSEFFLYSKLTVDKSRWQGNVPSCSV	447
Qy	655	THK 657	
Db	448	MHE 450	
RESULT 20			
US-09-825-012-55			
; Sequence 55, Application US/09825012			
; Patent No. US20020122798A1			
; GENERAL INFORMATION:			
; APPLICANT: Young, Robert			
; TITLE OF INVENTION: Compounds for Targeting			
; FILE REFERENCE: 43191-256808			
; CURRENT APPLICATION NUMBER: US/09/825,012			
; CURRENT FILING DATE: 2001-04-03			
; PRIOR APPLICATION NUMBER: US 60/237,159			
; PRIOR FILING DATE: 2000-10-02			
; PRIOR APPLICATION NUMBER: GB 0008049.9			
; PRIOR FILING DATE: 2000-04-03			
; NUMBER OF SEQ ID NOS: 102			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 55			
; LENGTH: 741			
; TYPE: PRT			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Humanised HMFGL heavy Chain - DNase I fusion			
US-09-825-012-55			

Query Match 22.4%; Score 789; DB 10; Length 741;
Best Local Similarity 31.5%; Pred. NO. 3e-27;
Matches 209; Conservative 65; Mismatches 151; Indels 238; Gaps 17;

Qy	1	EVLQQSGPOLVPGKASVKISCKASKGYSFTGYMHWVKSPKGKLEWIGRINPNNGVTLL	60
		: :	
Db	20	QVLQVSQCAEVKKPGASKYCKSKAGSYTSAYIEWROAPKGLWEGLLPGSNSRY	79
		: :	
Qy	61	NOKFKDKATILTVDKSSPTAYMELRSLTSDSAYVCARSTMITNYVM DYWGOGTSTVTSS	120
		: :	
Db	80	NEXFKGRVTVRTSTNTAYMELSSLRSEDATYYCARYDFAWFA - YWGOGTLTVTSS	137
		: :	
Qy	121	AKTTPPSVYPPLAFGSAAQTNSMYTLCGLVKGYFPEPVTWTNWSLSSGVHTFPVQLSD	180
		: :	
Db	138	ASTKGPVFPLAPSSKSTSGCTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSS	197
		: :	
Qy	181	-LYTLSSSVTPVPSTWPSEPTVCNVAHPASSTVDKKIKVIPROSGGSPSEKSEINEKDLRK	239
		: :	
Db	198	GLVSLSSVTPVPSSSLGTQYICNVNHKPSNTKDKRKVEPKSC - - - - -	240
		: :	
Qy	240	KSELQGTALGNLKOIYYNSKATTSSEKSADQFLTNTLLFKGFFTGHPWYNLDLLDGLST	299
		: :	
Db	241	-----DKHTTCPPCPAPELLGGPSVF--LFPPKP--KDTLM-----	272

Qy	300	AATSEYEGSSVDLYGAYYGQCAGGTPNKCTACMYGGVTLLHDNNRLTEEEKKVPINLWIDGK	359
Db	273	-----ISRTPEVT-CVVVDVS-----HEDPEVKFNWYVDG-	301
Qy	360	QTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGCLNSDSFGKGVQRGGLIVFHSSSEGSTV	419
Db	302	---VEVINAKTKPREQ-----YNSTY---RVYSVLTVLHQDWLNGK	337
Qy	420	SYDLFDAQQGPDYPLLRIRYDRNTTISSTLSISLYLTTTSIVMTQPTSLLYSAGDRVTI	479
Db	338	EY-----	339
Qy	480	TKKASQSVSNDVAWYQKPGQSPKLLISYTSRYAGVPDRFSGGYGCTDFTLTISSVQAE	539
Db	340	KCKYS-----NKALPAP--IEKTIKAKGQP-----	363
Qy	540	DAVYFCQDYNSPPTFGGKTKLEIKRADAAPTVSIPTSPSSSEQLTSGGASVVCFLNNFYP	599
Db	364	-----REPOQVYTLPPSRDELTKNQVSLTCLVKGFYP	394
Qy	600	KDINVKWKIDGSERON-----GVLNSWTDODSKDSTYSMSSTLTLLTKDEYERHNSVTC	654
Db	395	SDIAVEWESNGQPNKYKTPPVLDSDGSEFFLYSKLTVDKSRMQQGNVESC	447
Qy	655	THK 657	
Db	448	MHE 450	
RESULT 21			
US-09-859-053-28			
; Sequence 28, Application US/09859053			
; Patent No. US20020102658A1			
; GENERAL INFORMATION:			
; APPLICANT: Tsuji, Takashi			
; APPLICANT: Tezuka, Katsunari			
; APPLICANT: Hori, No. US20020102658A1uaki			
; TITLE OF INVENTION: HUMAN MONOCLONAL ANTIBODY AGAINST A			
; TITLE OF INVENTION: COSTIMULATORY SIGNAL TRANSDUCTION MOLECULE			
; TITLE OF INVENTION: PHARMACEUTICAL USE THEREOF			
; FILE REFERENCE: 06501-079001			
; CURRENT APPLICATION NUMBER: US/09/859,053			
; CURRENT FILING DATE: 2001-05-16			
; PRIOR APPLICATION NUMBER: JP 2001-99508			
; PRIOR FILING DATE: 2001-03-30			
; PRIOR APPLICATION NUMBER: JP 2000-147116			
; PRIOR FILING DATE: 2000-05-18			
; NUMBER OF SEQ ID NOS: 43			
; SOFTWARE: FastSeq for Windows version 4.0			
; SEQ ID NO 28			
; LENGTH: 470			
; TYPE: prt			
; ORGANISM: Homo sapiens			
US-09-859-053-28			

Query Match	22.3%	Score 785.5;	DB 10;	Length 470;
Best Local Similarity	29.6%	Pred. NO. 2.9e-27;		
Matches 201; Conservative	75;	Mismatches 136;	Indels 267;	Gaps 15;

[illegible]

Db 200 VLOSGSLSSVTVFPSSNFGTQTCYNDHKPSNTKVDKTVKRCCECPCCPAPPA 259
QY 224 GPSEKSEINEKDLKRKSELQGTALGNLKOIYYNSKAITSSKSAQDLTNTLLPKGFF 283
Db 260 GPS-----VFLPPRP-----KDTLMI----- 276
QY 284 TGPWYNLLVDLGSTAATSEYSGSVLDLYGAYYQYQAGTGNPKTACMYGGVTLHDNRR 343
Db 277 -----SRTPEVT-CVVVDVS----- 290
QY 344 LTEKKVPINLWIDGKOTVPIDIKVYTSKKEVTVVOELDLQARHYLHGKFGLYNSDSFGK 403
Db 291 -HEDPEVQFNMYDQ-----VEVHNAKTPRE----- 316
QY 404 VQRLIVFHSSEGSTVSYDLFLDAQOQPYDPTLLRLIYRDNNTTISSTLSLSLYLTTSVMT 463
Db 317 -----EQFNSTRVSVLTVVHQD----- 335
QY 464 QTPTSLLVASAGDRVTITCKASQSVSNDVAVTQKPGQSPKLLISYTSRYAGVDPDFSGS 523
Db 336 -----WLNKEYKC-----KYSNK 349
QY 524 GYGDFTLTSSVOAEDAAYVFCODYNSPPTFCGGTKLETKRADAPTYSIFPPSSEQL 583
Db 350 GLPAPIEKTISKTKGQ-----PREPQVYTLPPSREEM 381
QY 584 TSGCASVVCFLNNFYPKDINVKKIDGSERON-----GVLNSWTDQDSKSTYSMSSTLT 638
Db 382 TKNOVSTCLVKGYPSDIAVEWESNGOPENNYKTTTPMLDS-----DGSFFLYSKLT 434
QY 639 LTKDEYERHNSYTCETHK 657
Db 435 VDKSRWQGNVFCVSMHE 453

RESULT 22
US-09-995-693-2
; Sequence 2, Application US/09995693
; Patent No. US20020136721A1
; GENERAL INFORMATION:
; APPLICANT: Schwall, Ralph H.
; Tabor, Kelly H.
; TITLE OF INVENTION: Hepatocyte Growth Factor Receptor
; Antagonists and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/995,693
; FILING DATE: 29-Nov. US20020136721A1-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/952,235
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0938P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 230 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-995-693--2

Query Match 22.1%; Score 780; DB 10; Length 230;
Best Local Similarity 67.7%; Pred. No. 2.7e-27;
Matches 151; Conservative 28; Mismatches 38; Indels 6; Gaps 3;

QY 1 EVLQOQSGPDLVRPGASVKISCKASGYSTGYMHVWYKQSPGKGLGWIGRINPNNGVTLY 60
Db 1 ZVQLVQSGAEVYKPGASVKISCKASGYSTGYMHVWYKQSPGKGLGWIGRINPNNGVTLY 60
QY 61 NQKFKDKATLTVDKSSSTAYMELRSLTSEDSAVYICARSTMTNIV--MDYWGOGTSVTV 118
Db 61 NPNFKDKATLNVDRSSNTAYMLLSLTSADSAVYICA---TYGSYVSPLDYWGOGTSVTV 117
QY 119 SSARTTPPSVYPLAPGSAAGTNSMVTGLCLVKGYPFPEPTVTNNSGSLSSGVHTFPAVLQ 178
Db 118 SSASTKGPSVEPLAPSSKSTSGGTAALGCLVKGYPFPEPTVTNNSGSLSSGVHTFPAVLQ 177
QY 179 SD-LYTLSSSVTVPSSTWPSSTVTCNVAPASSTKVDKIVPR 220
Db 178 SSGLYSLSSVTVPSSTSLGTQTYICNVNHHKPSNTKVDKVEPK 220

RESULT 23
US-09-747-669-3
; Sequence 3, Application US/09747669
; Patent No. US20020122807A1
; GENERAL INFORMATION:
; APPLICANT: Dan, Michael D.
; APPLICANT: Saleh, Mansoor
; TITLE OF INVENTION: ANTIGEN BINDING FRAGMENTS, DESIGNATED
; TITLE OF INVENTION: 4B5 THAT SPECIFICALLY DETECT CANCER CELLS, NUCLEOTIDES
; TITLE OF INVENTION: ENCODING THE FRAGMENTS, AND USE THEREOF FOR THE PROPHYLAXIS
; FILE REFERENCE: 316082001001
; CURRENT APPLICATION NUMBER: US/09/747,669
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 09/111,286
; PRIOR FILING DATE: 1998-07-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 476
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic construct
US-09-747-669-3

Query Match 22.1%; Score 779.5; DB 10; Length 476;
Best Local Similarity 31.0%; Pred. No. 5.2e-27;
Matches 208; Conservative 69; Mismatches 150; Indels 243; Gaps 17;

QY 1 EVLQOQSGPDLVRPGASVKISCKASGYSTGYMHVWYKQSPGKGLGWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVYKPGASVKISCKASGYSTGYMHVWYKQSPGKGLGWIGRINPNNGVTLY 79
QY 61 NQKFKDKATLTVDKSSSTAYMELRSLTSEDSAVYICARSTMTNIVMDYWGOG 113
Db 80 AQKFGQRTVTRNTSIRTAIYMWELSGRLSEDTAVYFCARNADNVEMAAIYHYGHDVWGQG 139
QY 114 TSVTVSSAKTPPSVYPLAPGSAAGTNSMVTGLCLVKGYPFPEPTVTNNSGSLSSGVHTF 173
Db 140 TTVTVSSASTKGPSVFPPLAPSSKSTSGGTAALGCLVKGYPFPEPTVTNNSGSLSSGVHTF 199
QY 174 PAVLQSD-LYTLSSSVTVPSSTWPSSTVTCNVAPASSTKVDKIVPRDGGSGPSEKSEEI 232
Db 200 PAVLQSGLYSLSSVTVPSSTSLGTQTYICNVNHHKPSNTKVDKVEPKSC----- 249

QY 287 PWYNDLLVLDGTAATSEYEGSSVDLYGAYGYOCAGGTPNKACMYGCVTLHDNNRLTE 346
| : : : : : |
Db 275 PEVTCVWLDLG-----RE 287
QY 347 EKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKFLYNSDSFGKQVOR 406
| : : : : : |
Db 288 DPEVQISWFDGKE---VHTAKTQSR-----QQFNG--- 316
QY 407 GLIVPHSSSEGSVSYDLFDQAQGYPDTLRIYRDNTTISSTLSLSLYLTYTIVMTQTP 466
| : : : : : |
Db 317 -----TYR-----VVSVP 325
QY 467 TSLVYSAGDRVTITCKASQSVNDVAVYQKPGPKLLISYTSRYAGVPDRFSGSGVG 526
| : : : : : |
Db 326 -----TEHQDWLT-----G 334
QY 527 TDFTLTSSVQAEAAVYFCQDYNPPTFGGKTLEIKRADA---APTVISFPSSSEQL 583
| : : : : : |
Db 335 KEFKCRVNIH-----DLPSP-----IERTISKARGRAHKPSVYVLPSPKEL 376
QY 584 TSGG-ASVVCFLNFPKIDINVKWKIDGSEONGVLSNWDQDSKDYTSMSSTLTLTND 642
| : : : : : |
Db 377 SSSDVTISITCLIKDYPPDIDVWQSGQEQEPERKHRMTTPQLDEGSGFLYLSKLSVDKS 436
QY 643 EYERHNSYTCEATHKT 658
| : : : : : |
Db 437 RWQOGDPTCAVMHET 452
RESULT 26
US-09-828-995B-11
; Sequence 11, Application US/09828995B
; Patent No. US20020165135A1
; GENERAL INFORMATION:
; APPLICANT: Heska Corporation
; APPLICANT: McCall, Catherine A.
; APPLICANT: Tang, Liang A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATED TO CANINE IGF AND CANINE IL-13 R
; FILE REFERENCE: AL-7
; CURRENT APPLICATION NUMBER: US/09/828,995B
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,874
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/195,659
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 11
; LENGTH: 470
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-828-995B-11
Query Match 21.6%; Score 759.5; DB 9; Length 470;
Best Local Similarity 27.8%; Pred. No. 3.6e-26;
Matches 189; Conservative 87; Mismatches 135; Indels 269; Gaps 14;
QY 1 EVOLQSGDPLVKPGASVKISKASGYSGFTGYHHVHWKQSPCKGLEWTGRINPNNGVTLY 60
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 20 EVQLVESGDLVKPGGSLRLSCVASGFTFSDYGMGSRQSPGKGLQWAAVS-NRGDTYY 78
QY 61 NQKFKDKATLTVDKSTTAYMELRSITSEDSAVVYCARSTMTITNVV-MDYNGCGTSVTVS 119
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 79 ADAVKGRFTISRDNKNTLYLQMSLKAEDTHYHCVTGVWPRHRYGMDHWGNGTSLFVS 138
QY 120 SAKTTTPSYIPLAGSAAGTNSMVLGLCLVKGYPFPEVTVTNWNSGSLSSGVHTFPFVQLQS 179
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 139 SASTTAPSVFPLAPSGSTSGSTVALACLVSQYFPPEVTVSNWNSGSLTSGVHTFPFVQLQS 198
QY 180 D-LYTLSSSVTVPSWTSETVTCNVHPASTKVDKKI-----VPRDSGGP 225
| | | | | : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 199 SGLYSLSTVTVPPSSRWPFSETFTCNVHPASNTKVDKVPKPESTCKICISPCVPVPSLGGP 258

QY 226 SEKSEEINEKDLKRKSELQGTALGNLKOIYYNSKAITSSSEKSAQDFLTNTLLFKGFFTG 285
| : : : : : |
Db 259 SVFIFFPKPKDILR-----IPTT----- 276
QY 286 HPWYNDLLVLDGTAATSEYEGSSVDLYGAYGYOCAGGTPNKACMYGCVTLHDNNRLT 345
| : : : : : |
Db 277 -PEITCVWLDLG-----R 288
QY 346 EKKVPINLWIDGKQTTVPIDKVKTSKKEV-----TVQELDLQARHYLHGKFLYNSD 398
| : : : : : | : : : : : | : : : : : | : : : : : |
Db 289 EDEPQISWFDGKE---VHTAKTQPREQFNSTYRVVSVLPPIHQDWLTGK----- 337
QY 399 SFGGVQVORGLIVFHSSEGSVSYDLFDQAQGYPDTLRIYRDNTTISSTLSLSLYLTYT 458
| : : : : : | : : : : : | : : : : : |
Db 338 EFKCRVNIH-----HGLSPPIERTISKARGQ----- 362
QY 459 SIWMTQTPTSLVYSAGDRVTITCKASQSVNDVAVYQKPGSPKLLISYTSRYAGVPD 518
| : : : : : |
Db 363 -----AHQ----- 365
QY 519 RPSGSGYGTDFTLTSSVQAEAAVYFCQDYNPPTFGGKTLEIKRADAAPTVISIIPP 578
| : : : : : | : : : : : |
Db 366 -----PSVVVLP 373
QY 579 SSEQLTSGG-ASVVCFLNFPKIDINVKWKIDGSEONGVLSNWDQDSKDYTSMSSTL 637
| : : : : : | : : : : : | : : : : : | : : : : : | : : : : : |
Db 374 SPKELSSDVTTLTCLIKDFPPEIDVWQSGQEPESKYHTTAPQLDEGSGFLYLSKL 433
QY 638 TLTKDEYERHNSYTCEATHK 657
| : : : : : | : : : : : |
Db 434 SVDKSRWQOGDPTCAVMHE 453
RESULT 27
US-09-811-384-12
; Sequence 12, Application US/09811384
; Patent No. US20020081294A1
; GENERAL INFORMATION:
; APPLICANT: Bednar, Martin M.
; APPLICANT: Thomas, G. Roger
; APPLICANT: Gross, Cordell E.
; TITLE OF INVENTION: ANTI-CD18 ANTIBODIES IN STROKE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/811,384
; FILING DATE: 20-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/251652
; FILING DATE: 17-FEB-2000
; APPLICATION NUMBER: 08/788800
; FILING DATE: 22-JAN-1997
; APPLICATION NUMBER: 60/093038
; FILING DATE: 23-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: PI1729C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530


```

; PRIOR FILING DATE: 2000-08-29
; PRIOR APPLICATION NUMBER: 60/152,248
; PRIOR FILING DATE: 1999-09-03
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-514-8

Query Match          21.0%   Score 740:   DB 10:   Length 240:
Best Local Similarity 61.3%   Pred. No. 1.4e-25;
Matches 146; Conservative 32; Mismatches 50; Indels 10; Gaps 2:

QY  444  ISSTSLSISLYLTT----SIVMTQTPTSLVVSAGDRVTITCKASQSV-----SNDVAV 493
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  2  VLOQVFLSLWLTSGAYGDIVMTQSPDSLAVSLGERATINCKSSQTVLYSSDNKNYLAW 61

QY  494  YQKPGQSPKLLISYTSRYAGVDPDRFGSGYGTDFTLTISVQAEADAAYFCQDYNSP 553
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  62  YQKPGQPPKLLIYWASTRESGVDPDRFGSGSGTDFTLTISLQAEDEVAVYCCQYVSTP 121

QY  554  PTFEGGQKLEIKRADAPTIVSIFPPSSQLTSGASVVCFLNNFYPKDINVKWKIDGSR 613
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  122  YSFGQGQKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCCLLNNFYPREAKVQWKVDNALQ 181

QY  614  QNGVLNSWTDQDSKDSYSMSSTLTLTKDEYERHNSYTCETHKTSTSPIVKSFNRR 671
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
DB  182  SGNQSESVTEQDSKDSYSLSTLTLSKADYEHKHYACEVTHQGLSSPVTKSFNRGE 239

RESULT 29
US-09-940-166A-6
; Sequence 6, Application US/09940166A
; Patent No. US2002005832A1
; GENERAL INFORMATION:
; APPLICANT: Blank, Gregory S.
; Narindray, Daljit S.
; Zapata, Gerardo A.
; TITLE OF INVENTION: Protein Recovery
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,166A
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/097,309
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwartz, Timothy R.
; REGISTRATION NUMBER: 32171
; REFERENCE/DOCKET NUMBER: P1105R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-7467
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear

```



```
Db 62 FSGSGDYYTLTSSLOPEDATYVCOGNLTLPFTFGQGTKEIKRTVAAPSVFIPTPS 121
QY 580 SEQLTSGGASVVCFLNNFKPKDINVKWIDGSRQNGVLNSWTDODSKDSTYSMSLTTL 639
Db 122 DEQLKSGTASVCLLNNFVPRAKVQWVDNALQSGNSQESVTEQDSKDSYLSLTTL 181
QY 640 TKDEYERHNSYTCETHAKTSTSPVSKFNRNE 671
Db 182 SKADYEKKHVVACEVTHQGLSSPVTKSFNRGE 213

RESULT 32
US-09-995-693-1
; Sequence 1, Application US/09995693
; Patent No. US20020136721A1
; GENERAL INFORMATION:
; APPLICANT: Schwall, Ralph H.
; TITLE OF INVENTION: Hepatocyte Growth Factor Receptor
; Antagonists and Uses Thereof
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/995,693
; FILING DATE: 29-NOV-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/952,235
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Marschang, Diane L.
; REGISTRATION NUMBER: 35,600
; REFERENCE/DOCKET NUMBER: P0938P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5416
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 220 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-995-693-1
Query Match 20.6%; Score 726; DB 10; Length 220;
Best Local Similarity 63.3%; Pred. No. 5, 1e-25;
Matches 138; Conservative 34; Mismatches 40; Indels 6; Gaps 1;

QY 460 IYVMTPTSLVYAGRVITITCKASQSV-----SNDVAYOQKPGQSPKLLISYTSRY 513
Db 2 IMWSQSPSLTVSGEKVTVCSKSSQSLLYTSSQKNYLAWYQKPGQSPKLLIYNASTRE 61
QY 514 AGVPRFSGSGYCTDFTLTITSSVQAEADAVYFCQDYNPPFTFGGTKLEIKRAADAPT 573
Db 62 SGVPRFTGSGSGTDFLTITSVKADDLAVYVYCOQYAYPWTFTGGTKLEIKRTVAAPSV 121
QY 574 SIFFPSSEQLTSGGASVVCFLNNFKPKDINVKWIDGSRQNGVLNSWTDQDSKDSYMS 633
Db 122 FIFPSDEQLKSGTASVCLLNNFVPRAKVQWVDNALQSGNSQESVTEQDSKDSYSL 181
QY 634 SSTLTITKDEYERHNSYTCETHAKTSTSPVSKFNRNE 671
```

```
Db 182 SSTLTSLKADYEKKHVVACEVTHQGLSSPVTKSFNRGE 219

RESULT 33
US-09-940-166A-7
; Sequence 7, Application US/09940166A
; Patent No. US20020058324A1
; GENERAL INFORMATION:
; APPLICANT: Blank, Gregory S.
; Narindray, Daljit S.
; Zapata, Gerardo A.
; TITLE OF INVENTION: Protein Recovery
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,166A
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/097,309
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwartz, Timothy R.
; REGISTRATION NUMBER: 32171
; REFERENCE/DOCKET NUMBER: P1105R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-7467
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 300 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-940-166A-7
Query Match 20.6%; Score 725.5; DB 10; Length 300;
Best Local Similarity 53.5%; Pred. No. 7e-25;
Matches 147; Conservative 43; Mismatches 62; Indels 23; Gaps 5;

QY 1 EVLOQSGPDLVKGASVKISKASGYSTGYIMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db 24 EVQLVESGGGLVQPGGSLRLSCATGYTETMTMMRWQAPGKLEWAGINPKNGGTS 83
QY 61 NQKFKDKATLTVDKSTTAYMELSLTSDSAVYICARSTMTITNY-----VMDYNGQGT 115
Db 84 NORFMDRFTISVDKSTSTAYMOMNSLRADTAVYICARWRGL-NYGFVRYFDVWQGT 142
QY 116 VTVSSAKTTPPSVYPLAPGSAAGTNSMTVLGCLVKGYPPEVTVTNWNSGLSSGVHTT 175
Db 143 VTVSSASTKGPVFPPLAPSSKSTSGTAAALGCLVXDYFPPEVTVSNWNSGALTSGV 202
QY 176 VLQSD-LYTLSSVTVPSSTWPSSETVTCNVAHPASTKVDKKIVPRDS----- 222
Db 203 VLQSSGLYSLSSVTVPPSSSLGTQTVICNVNHPKSNVTKVDKKVEPKSCDKTHTCP 262
QY 223 ---GGPSEKSEIKEDLKKKSELQGTALGNLKOI 254
Db 263 ELLGGRMKOLEDKVELLSKNYHLE-NEVARUKKL 296
```

```
RESULT 34
US-09-825-012-66
; Sequence 66, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825.012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66
; LENGTH: 515
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 Fd - DNase I fusion
US-09-825-012-66

Query Match      20.5%; Score 722.5; DB 10; Length 515;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVQLQSGPDLVKPCASVKISKASGYSTGYMHVWVKQSPGKGLEWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPGASVKVKSCASGYTFSAYIEWVRQAPGKGLEWYGEILPGSNNSRY 79
;
Qy 61 NQKFKDKATLTVDKSSSTAYMELSLTSEDSAVYYCARSTMITNVMYDYGQGTSTVTVSS 120
Db 80 NEKFKGRVTVTRDTSTNTAYMELSLRSEDYAVYYCARSDYFAWFA--YWGQGTLLTVTVSS 137
;
Qy 121 AKTTPPSVYPLAPGSAAGTNSMVTGLCLVKGYFPEPVTVTVWNSGSLSSGVTFFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS 197
;
Qy 181 -LYTLSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIIVPR 220
Db 198 GLYSLSVVTVPSSSLGTQTYICNVNHHKPSNTKVDKKVEPK 238
;
RESULT 36
US-09-825-012-76
; Sequence 76, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825.012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 76
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFG1 Fd - DNase I fusion
US-09-825-012-76

Query Match      20.5%; Score 722.5; DB 10; Length 519;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVQLQSGPDLVKPCASVKISKASGYSTGYMHVWVKQSPGKGLEWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPGASVKVKSCASGYTFSAYIEWVRQAPGKGLEWYGEILPGSNNSRY 79
;
Qy 61 NQKFKDKATLTVDKSSSTAYMELSLTSEDSAVYYCARSTMITNVMYDYGQGTSTVTVSS 120
Db 80 NEKFKGRVTVTRDTSTNTAYMELSLRSEDYAVYYCARSDYFAWFA--YWGQGTLLTVTVSS 137
;
Qy 121 AKTTPPSVYPLAPGSAAGTNSMVTGLCLVKGYFPEPVTVTVWNSGSLSSGVTFFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS 197
;
Qy 181 -LYTLSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIIVPR 220
Db 198 GLYSLSVVTVPSSSLGTQTYICNVNHHKPSNTKVDKKVEPK 238
;
RESULT 37
US-09-825-012-80
; Sequence 80, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825.012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
```

; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 80
; LENGTH: 519
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMF1 Fd - DNase I fusion
US-09-825-012-80

Query Match 20.5%; Score 722.5; DB 10; Length 519;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVLOQSGPDLVKPGASVKISKASGYSTGYMHVWQSPGKLEWIGRIINPNNGVTLY 60
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Db 20 QVQLVQSGAEVKKPGASVKVSKASGYTFSAYIEWRQAPGKLEWGEILPGSNNSRY 79
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||

Qy 61 NQFKDKATLTVDKSTTAYMELSLTSDSAVYYCARSTMTITNYMDYWGOGTSVTYSS 120
: ||| : ||: ||: ||||| || |||:||||| : ||||| |||||
Db 80 NEKFGRTVTTRDTSTNTAYMELSLRSEDYAVYYCARSDYDFAWFA--YWGQGLTVTVSS 137
: ||| : ||: ||: ||||| || |||:||||| : ||||| |||||

Qy 121 AKTTPSVYPLAPGSAAGTNSMVTGLCLVKGYFPEPVTVWNSGSLSGVHTFPVAVLQSD 180
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Db 138 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSS 197
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||

Qy 181 -LYTLSSSVTPSPSTWPSSTVTCNVNVAHPASSTKVDKKIVPR 220
: ||| ||| |||||: ||: ||| || |||: |||||: ||
Db 198 GLYSLSVVTPSPSSLGQTGYICNVNHNKPSNTKVDKKVEPK 238
: ||| ||| |||||: ||: ||| || |||: |||||: ||

RESULT 38
US-09-825-012-71
; Sequence 71, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMF1 Fd - DNase I fusion
US-09-825-012-71

Query Match 20.5%; Score 722.5; DB 10; Length 521;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVLOQSGPDLVKPGASVKISKASGYSTGYMHVWQSPGKLEWIGRIINPNNGVTLY 60
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Db 20 QVQLVQSGAEVKKPGASVKVSKASGYTFSAYIEWRQAPGKLEWGEILPGSNNSRY 79
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||

Qy 61 NQFKDKATLTVDKSTTAYMELSLTSDSAVYYCARSTMTITNYMDYWGOGTSVTYSS 120
: ||| : ||: ||: ||||| || |||:||||| : ||||| |||||
Db 80 NEKFGRTVTTRDTSTNTAYMELSLRSEDYAVYYCARSDYDFAWFA--YWGQGLTVTVSS 137
: ||| : ||: ||: ||||| || |||:||||| : ||||| |||||

Qy 121 AKTTPSVYPLAPGSAAGTNSMVTGLCLVKGYFPEPVTVWNSGSLSGVHTFPVAVLQSD 180
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Db 138 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSS 197
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||

Qy 181 -LYTLSSSVTPSPSTWPSSTVTCNVNVAHPASSTKVDKKIVPR 220
: ||| ||| |||||: ||: ||| || |||: |||||: ||
Db 198 GLYSLSVVTPSPSSLGQTGYICNVNHNKPSNTKVDKKVEPK 238
: ||| ||| |||||: ||: ||| || |||: |||||: ||

RESULT 39
US-09-825-012-85
; Sequence 85, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 85
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMF1 Fd - DNase I fusion
US-09-825-012-85

Query Match 20.5%; Score 722.5; DB 10; Length 525;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVLOQSGPDLVKPGASVKISKASGYSTGYMHVWQSPGKLEWIGRIINPNNGVTLY 60
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Db 20 QVQLVQSGAEVKKPGASVKVSKASGYTFSAYIEWRQAPGKLEWGEILPGSNNSRY 79
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||

Qy 61 NQFKDKATLTVDKSTTAYMELSLTSDSAVYYCARSTMTITNYMDYWGOGTSVTYSS 120
: ||| : ||: ||: ||||| || |||:||||| : ||||| |||||
Db 80 NEKFGRTVTTRDTSTNTAYMELSLRSEDYAVYYCARSDYDFAWFA--YWGQGLTVTVSS 137
: ||| : ||: ||: ||||| || |||:||||| : ||||| |||||

Qy 121 AKTTPSVYPLAPGSAAGTNSMVTGLCLVKGYFPEPVTVWNSGSLSGVHTFPVAVLQSD 180
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||
Db 138 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSS 197
: ||| ||| : |||||:|||||: ||: ||:|||||: ||: ||: ||

Qy 181 -LYTLSSSVTPSPSTWPSSTVTCNVNVAHPASSTKVDKKIVPR 220
: ||| ||| |||||: ||: ||| || |||: |||||: ||
Db 198 GLYSLSVVTPSPSSLGQTGYICNVNHNKPSNTKVDKKVEPK 238
: ||| ||| |||||: ||: ||| || |||: |||||: ||

RESULT 40
US-09-825-012-43
; Sequence 43, Application US/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 43
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMF1 Fd - DNase I fusion

US-09-825-012-43

Query Match 20.5%; Score 722.5; DB 10; Length 527;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;
Qy 1 EVQLQSGPDLVKPGASVKISKASGYSTGYMHVWQSPGKGLGWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPKASVKVSKASGYTFSAYIEWVRQAPGKGLGWGEILPGSNNSRY 79
Qy 61 NQKFKDKATLTVDKSSTTAYMELSLRSTSDSAVYYCARSTMTITNYMDYWGOGTSVTVSS 120
Db 80 NEKFKGRVTVTTRDTSTNTAYMELSLRSEDATVYYCARSYDFAWFA--YWGQGLTVTVSS 137
Qy 121 AKTTPPSVYPLAPGSAAGTNSMTLGLCLVKGYFPEPVTVTVNSGSLSSGVHFFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS 197
Qy 181 -LYTLSSSVTPSPSWPSETVTCNVAHPASSTKVDKIVPR 220
Db 198 GLYSLSVVTPSSSLGTQTYICNVNHNKPSNTKVDKVEPK 238

RESULT 41

US-09-825-012-95
; Sequence 95, Application us/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:
; APPLICANT: Young, Robert
; TITLE OF INVENTION: Compounds for Targeting
; FILE REFERENCE: 43191-256808
; CURRENT APPLICATION NUMBER: US/09/825,012
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/237,159
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: GB 0008049.9
; PRIOR FILING DATE: 2000-04-03
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 95
; LENGTH: 529
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Humanised HMFGL Fd - DNase I fusion
US-09-825-012-95
Query Match 20.5%; Score 722.5; DB 10; Length 529;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;
Qy 1 EVQLQSGPDLVKPGASVKISKASGYSTGYMHVWQSPGKGLGWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPKASVKVSKASGYTFSAYIEWVRQAPGKGLGWGEILPGSNNSRY 79
Qy 61 NQKFKDKATLTVDKSSTTAYMELSLRSTSDSAVYYCARSTMTITNYMDYWGOGTSVTVSS 120
Db 80 NEKFKGRVTVTTRDTSTNTAYMELSLRSEDATVYYCARSYDFAWFA--YWGQGLTVTVSS 137
Qy 121 AKTTPPSVYPLAPGSAAGTNSMTLGLCLVKGYFPEPVTVTVNSGSLSSGVHFFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS 197
Qy 181 -LYTLSSSVTPSPSWPSETVTCNVAHPASSTKVDKIVPR 220
Db 198 GLYSLSVVTPSSSLGTQTYICNVNHNKPSNTKVDKVEPK 238

RESULT 42

US-09-825-012-90
; Sequence 90, Application us/09825012
; Patent No. US20020122798A1
; GENERAL INFORMATION:

APPLICANT: Young, Robert
TITLE OF INVENTION: Compounds for Targeting
FILE REFERENCE: 43191-256808
CURRENT APPLICATION NUMBER: US/09/825,012
CURRENT FILING DATE: 2001-04-03
PRIOR APPLICATION NUMBER: US 60/237,159
PRIOR FILING DATE: 2000-10-02
PRIOR APPLICATION NUMBER: GB 0008049.9
PRIOR FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 102
SOFTWARE: PatentIn version 3.1
SEQ ID NO 90
LENGTH: 531
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Humanised HMFGL Fd - DNase I fusion
US-09-825-012-90

Query Match 20.5%; Score 722.5; DB 10; Length 531;
Best Local Similarity 62.9%; Pred. No. 1.5e-24;
Matches 139; Conservative 32; Mismatches 47; Indels 3; Gaps 2;

Qy 1 EVQLQSGPDLVKPGASVKISKASGYSTGYMHVWQSPGKGLGWIGRINPNNGVTLY 60
Db 20 QVQLVQSGAEVKKPKASVKVSKASGYTFSAYIEWVRQAPGKGLGWGEILPGSNNSRY 79
Qy 61 NQKFKDKATLTVDKSSTTAYMELSLRSTSDSAVYYCARSTMTITNYMDYWGOGTSVTVSS 120
Db 80 NEKFKGRVTVTTRDTSTNTAYMELSLRSEDATVYYCARSYDFAWFA--YWGQGLTVTVSS 137
Qy 121 AKTTPPSVYPLAPGSAAGTNSMTLGLCLVKGYFPEPVTVTVNSGSLSSGVHFFPAVLQSD 180
Db 138 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS 197
Qy 181 -LYTLSSSVTPSPSWPSETVTCNVAHPASSTKVDKIVPR 220
Db 198 GLYSLSVVTPSSSLGTQTYICNVNHNKPSNTKVDKVEPK 238

RESULT 43

US-09-811-384-10
; Sequence 10, Application us/09811384
; Patent No. US20020081294A1
; GENERAL INFORMATION:
; APPLICANT: Bednar, Martin M.
; Thomas, G. Roger
; Gross, Cordell E.
; TITLE OF INVENTION: ANTI-CD18 ANTIBODIES IN STROKE
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/811,384
; FILING DATE: 20-Dec-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/251652
; FILING DATE: 17-FEB-2000
; APPLICATION NUMBER: 08/788800
; FILING DATE: 22-JAN-1997
; APPLICATION NUMBER: 60/093038
; FILING DATE: 23-JAN-1996

```
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: PL729C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 232 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-811-384-10

Query Match          20.4%; Score 718.5; DB 10; Length 232;
Best Local Similarity 61.1%; Pred. No. 1.1e-24;
Matches 138; Conservative 35; Mismatches 46; Indels 7; Gaps 3;

QY 1 EVQLQQSGPDLVKGASVKISKASGYSTGYMHVWKQSPGKLEWIGRINPNNGVTLY 60
   |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 EVQLVESGGGLVQPGGSLRLSCATSGYTFETETWMMRQAPGKLEWVAGINPKNGTSH 60

QY 61 NQKFKDKATLTVDKSSITAYMELRSLTSEDSAVYYCARSTMITNY-----VMDYWGOGTS 115
   |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 61 NQRFMDRFTISVDKSTSTAYMQMNSLRAEDTAVYYCARWRGL-NYGFDRVYFDVWGQGT 119

QY 116 VTVSSAKTTPPSVYPLAPGSAATNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPA 175
   |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 120 VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSNWNSGALTSGVHTFPA 179

QY 176 VLQSD-LYTLSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIIVPR 220
   |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 180 VLQSSGLYSLSSVTVPSSTWPSSTGTTQTYICNVNHPKPSNTKVDKKVEPK 225

RESULT 44
US-09-940-166A-1
; Sequence 1, Application US/09940166A
; Patent No. US2002058324A1
; GENERAL INFORMATION:
; APPLICANT: Blank, Gregory S.
; Narindray, Daljit S.
; Zapata, Gerardo A.
; TITLE OF INVENTION: Protein Recovery
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/940,166A
; FILING DATE: 27-Aug-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/097,309
; FILING DATE: 13-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Schwartz, Timothy R.
; REGISTRATION NUMBER: 32171
; REFERENCE/DOCKET NUMBER: PL105R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-7467
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-940-166A-1
```

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-940-166A-1

Query Match          20.4%; Score 718.5; DB 10; Length 241;
Best Local Similarity 61.1%; Pred. No. 1.1e-24;
Matches 138; Conservative 35; Mismatches 46; Indels 7; Gaps 3;

QY 1 EVQLQQSGPDLVKGASVKISKASGYSTGYMHVWKQSPGKLEWIGRINPNNGVTLY 60
   |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 EVQLVESGGGLVQPGGSLRLSCATSGYTFETETWMMRQAPGKLEWVAGINPKNGTSH 60

QY 61 NQKFKDKATLTVDKSSITAYMELRSLTSEDSAVYYCARSTMITNY-----VMDYWGOGTS 115
   |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 61 NQRFMDRFTISVDKSTSTAYMQMNSLRAEDTAVYYCARWRGL-NYGFDRVYFDVWGQGT 119

QY 116 VTVSSAKTTPPSVYPLAPGSAATNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPA 175
   |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 120 VTVSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSNWNSGALTSGVHTFPA 179

QY 176 VLQSD-LYTLSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIIVPR 220
   |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 180 VLQSSGLYSLSSVTVPSSTWPSSTGTTQTYICNVNHPKPSNTKVDKKVEPK 225

RESULT 45
US-09-736-371B-21
; Sequence 21, Application US/09736371B
; Patent No. US20020131968A1
; GENERAL INFORMATION:
; APPLICANT: Waldmann, Herman
; Applicant: Frewin, Mark
; TITLE OF INVENTION: AGLYCOSYLATED ANTIBODIES
; FILE REFERENCE: Waldmann
; CURRENT APPLICATION NUMBER: US/09/736,371B
; CURRENT FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 9815909.8
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: PCT/GB99/02380
; PRIOR FILING DATE: 1999-07-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 449
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-371B-21

Query Match          20.4%; Score 718.5; DB 10; Length 449;
Best Local Similarity 29.3%; Pred. No. 1.9e-24;
Matches 194; Conservative 71; Mismatches 161; Indels 237; Gaps 17;

QY 1 EVQLQQSGPDLVKGASVKISKASGYSTGYMHVWKQSPGKLEWIGRINPNNGVTLY 60
   |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 1 EVQLLESGGGLVQPGGSLRLSCAASGFTSFSPFPAWVQAPGKLEWVSTLSTSGRTYY 60

QY 61 NQKFKDKATLTVDKSSITAYMELRSLTSEDSAVYYCARSTMITNYMDYWGOGTSVTVSS 120
   |||| || |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 61 RDSVKGRFTISRDNKNTLYLQMNSLRAEDTAVYYCAKFRQYSG-GFDYWGQGTLVTVSS 119

QY 121 AKTTPPSVYPLAPGSAATNSMVTLGCLVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
   |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 120 ASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSNWNSGALTSGVHTFPAVLQSS 179

QY 181 -LYTLSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKKIIVPRDSSGSPSEKSEINEKDLRK 239
   |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
Db 180 GLYSLSVTVPSSTGTTQTYICNVNHPKPSNTKVDKKVEPKSC----- 222

QY 240 KSELQGTALGNLKOIYYNYSKAITSSSEKSDQFLTNTLLFKGFTGHPWYNLLVLQGST 299
   |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| |||| ||||
```

Db 223 -----DKTHTPPCPAPELIGGPSVF--LFPKPK--KDTLM----- 254
Qy 300 AATSEYEGSSVDLYGAYGQCAGGTNKTACMYGGVTLHDNNRLTEKKVYPINLWIDGK 359
Db 255 -----ISRTEPT-CVVVDVS-----HEDPEVKFNWYDVG- 283
Qy 360 QTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDFGKQVQGLIVFHSSEGSTV 419
Db 284 --VEVHNKTKPRE----- 295
Qy 420 SYDLFDAQGQVPTDLLRIYRDNTTISSTLSISLYLTTTSIVMTQPTTSLLSVAGDRVTI 479
Db 296 -----EQYAST-YRWVSVLTVLHODWLNKEY----- 321
Qy 480 TKASQSVSNDVAVYQOKPGQSKLLISYTSRSGVAGVDRFSGSGYGTDTTLTSSVQAE 539
Db 322 CKKVS-----NKALPAP---TEKTSKAKGP----- 345
Qy 540 DAAVYFCQDYNSPPTFGGTGKLEIKRADAAPTYSIFSPSEQLTSGGASVVCFLNNFYP 599
Db 346 -----REPQYITLPPSRDELTKNOVSLTCLVKGFYP 376
Qy 600 KQINVKWIDGSEKQ-----GVLSNWDQDQSKDSTYSMSSTLTLTKDEYERHNSYTCFA 654
Db 377 SDIAVENESNGQPNKYKTTTPVLDSE-----DCSFFLYSKLTVDKSRWQOQGNVFCSV 429
Qy 655 THK 657
Db 430 MHE 432

RESULT 46
US-09-828-995B-17
; Sequence 17, Application US/09828995B
; Patent No. US20020165135A1
; GENERAL INFORMATION:
; APPLICANT: Heskia Corporation
; APPLICANT: McCall, Catherine A.
; APPLICANT: Tang, Liang A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS RELATED TO CANINE IgG AND CANINE IL-13
; FILE REFERENCE: AL-7
; CURRENT APPLICATION NUMBER: US/09/828,995B
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 60/195,874
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/195,659
; PRIOR FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 474
; TYPE: PRT
; ORGANISM: Canis familiaris
; US-09-828-995B-17

Query Match 20.4%; Score 717.5; DB 9; Length 474;
Best Local Similarity 27.4%; Pred. No. 2.2e-24;
Matches 186; Conservative 80; Mismatches 149; Indels 265; Gaps 13;

Qy 1 EVQLQQSGDPLVKPGASVKISKASGYSFTGYMHWVKQSPGKLEWIGRIINPNNGVTLY 60
Db 20 DVQLVESGDDLKPGGSLRLSCVAGSFTFSSCAMSWVRQSPGKPGQWVATIRYDGSIIY 79
Qy 61 NQKFKDQATLVKQSSITAYMELRSLTSEDSAVYICARSTMIT--NIVMDYWGOGTSVTV 118
Db 80 ADAYKGRFSTRNAKNTVYLQMSLRAEDTAVYICAKAPPDYSHYGMVWPGTSLFV 139
Qy 119 SSAKTTPEPVYPLAPGSAATQVNLGLCKVGFPEPTVTNWSGLSSGVHTFPAVLQ 178
Db 140 SSASTTAPVFPPLAPSGCGSGSTVALACLVSIGYIPEVTIVSNWSVLTSGVHTFPPVLQ 199
Qy 179 SD-LYTLSSSYVTPSSTWPSSTVTCNVAHPASSTKVDKIVPRDS----- 222

Db 200 SSGLYSLSMVTPSSRWPSETFTCNVAHPATNTKVDKPVAKCEKCNMCNCPGCGL 259
Qy 223 -GGPSEKSEEEINEKDLRKKSELOGTALGNLQKIYYYSKAITSEKSDAQPLTWTLFPKG 281
Db 260 LGGPS----- 264
Qy 282 FFTGHPWYNDDLVDLSTAATAATSEYEGSSVDLYGAYGQCAGGTNKTACMYGGVTLHDN 341
Db 265 VFIFPPKPKDLV---TARTPTVTCVVVDLD----- 292
Qy 342 NRUTEKKKVINLWIDGKQ-----TTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNS 397
Db 293 ---PENPEVQISWFDVKQVQTANTQPREQSNCTYRV-VSVLPIGHQDWLSGK----- 342
Qy 398 DSGGKQVQGLIVFHSSEGSTVSVDLFDACQOYPTDLLRIYRDNTTISSTLSLSLYLT 457
Db 343 -QFKCKV-----NNKALPS----- 355
Qy 458 TSIVMTQPTTSLLSVAGDRVTITCKASQSVSNDVAVYQOKPGQSKLLISYTSRSGVAGV 517
Db 356 -----PIEELIS-----KTPGQAHQ----- 370
Qy 518 DRFSGSGYGTDTTLTSSVQAEAAVYFCQDYNSPPTFGGTGKLEIKRADAAPTVSIFP 577
Db 371 -----PNVYVLP 377
Qy 578 PSSEQLTSGGASVVCFLNNFYPKIDINVKWIDGSEKQGVLSNWDQDQSKDSTYSMSSTL 637
Db 378 PSRDEMSKNTVTLTCLVKDFPFPEIDVEMOSNGOQEPESKYRMTPPQLDEGVSFLYSKL 437
Qy 638 TLTKDEYERHNSYTCETHK 657
Db 438 SVDKSRWQRGDTFICAVMHE 457

RESULT 47
US-10-011-125-5
; Sequence 5, Application US/10011125
; Patent No. US20020142388A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Christina Yu-Ching
; TITLE OF INVENTION: BACTERIAL HOST STRAINS
; FILE REFERENCE: P1804R1
; CURRENT APPLICATION NUMBER: US/10/011,125
; CURRENT FILING DATE: 2001-12-07
; PRIOR APPLICATION NUMBER: US 60/256,162
; PRIOR FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 5
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence is synthesized.
; US-10-011-125-5

Query Match 20.3%; Score 716; DB 12; Length 212;
Best Local Similarity 64.2%; Pred. No. 1.3e-24;
Matches 136; Conservative 29; Mismatches 45; Indels 2; Gaps 1;

Qy 460 IVMTQPTTSLLSVAGDRVTITCKASQSVSNDVAVYQOKPGQSKLLISYTSRSGVAGVDR 519
Db 2 IQMTQSPSSLSASVGDRVTITCRASQDINLYNLYQKPGKAPKLLIYYS--HSGVPSR 59
Qy 520 FSGSGYGTDTTLTSSVQAEAAVYFCQDYNSPPTFGGTGKLEIKRADAAPTVSIFPPS 579
Db 60 FSGSGSGTDTTLTSSLPQEDFATYYCQOQNTLPTFGGTGKVEIKRTVAAPSVFIPPS 119
Qy 580 SEQLTSGGASVVCFLNNFYPKIDINVKWIDGSEKQGVLSNWDQDQSKDSTYSMSSTLTL 639
Db 120 DEQLKSGTASVAVCLLNNFYPREAKVQWMDNALSGNSQESVTEODSKDSTYSLSSTLTL 179
Qy 640 TKDEYERHNSYTCETHKSTSPIVKSFNRNE 671


```

Db      180 SKADYEKKHYACEVTHQGLSSPVTKSFNRGE 211
       :|::||: | || |: ::|: ||||| |
       :|::||: | || |: ::|: ||||| |

RESULT 48
US-10-006-771A-6
; Sequence 6, Application US/10006771A
; Patent No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006,771A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/250,090
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens and Mus sp.
US-10-006-771A-6

Query Match          20.3%; Score 715.5; DB 9; Length 232;
Best Local Similarity 61.7%; Pred. No. 1.5e-24;
Matches 140; Conservative 30; Mismatches 50; Indels 7; Gaps 2;

QY   451 ISLYLYTTS-----IWMQTPTSLLSAGDRVITITCKASQSVDVAWYQQRPQSPKL 504
Db   6 IILFLVATATGVHSDIQLTQSPSLSASVGDRVTITCKAQDVGTSVAWYQQKPKAPKL 65
QY   505 LISYTSRYAGVPDRFSGSGYGDTFTLTISSVOAEDAAYFCQODYNSPPTFGGGTKLEI 564
Db   66 ILVYSTRHTGVSRFSFGSGGTDFTFTISSLPEDLATYYCQ-YSLYSRFGGGTKVEI 124
QY   565 KRDAAPTIVSIFPPSSQLSTGSAGVCVFLLNFPKDINVKWKIDGSEKNGVLNSWTDQ 624
Db   125 KRTVAAPSIFIPPSDQLKSSTASVYCLLNFPYREAKVQWKVDNALQSGNSQESVTEQ 184
QY   625 DSKDSTYSMSSTLTLTKDEVERHNSYTCEATHKTSTSPIVKSFNRRNE 671
Db   185 DSKDSTYLSLSTLTLSKADYEKKHYACEVTHQGLSSPVTKSFNRGE 231

RESULT 49
US-09-249-011A-22
; Sequence 22, Application US/09249011A
; Patent No. US20020176855A1
; GENERAL INFORMATION:
; APPLICANT: CO, MAN SUNG
; APPLICANT: VASQUEZ, MAXIMILIANO
; APPLICANT: CARENNO, BEATRIZ
; APPLICANT: CELNIKER, ABBIE CHERYL
; APPLICANT: COLLINS, MARY
; APPLICANT: GOLDMAN, SAMUEL
; APPLICANT: GRAY, GARY S.
; APPLICANT: KNIGHT, ANDREA
; APPLICANT: O'HARA, DENISE
; APPLICANT: RUP, BONITA
; APPLICANT: VELDMAN, GEERTRUIDA M.
; TITLE OF INVENTION: HUMANIZED IMMUNOGLOBULIN REACTIVE WITH B7-2 AND METHODS
; FILE REFERENCE: 08702.0081-000000
; CURRENT APPLICATION NUMBER: US/09/249, 011A
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 22
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-249-011A-22

```

Search completed: January 8, 2003, 12:05:26
Job time : 25.6232 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:55:11 ; Search time 25.2464 Seconds
(without alignments)
2558.872 Million cell updates/sec

Title: US-09-900-766-1
Perfect score: 3522
Sequence: 1 EVQLQQSGPDLVKPEASVKI.....EATHKTSTSPIVKSFNRES 672

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 9613422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1107	31.4	257	A28179	enterotoxin E prec
2	1036.5	29.4	225	S37484	Ig kappa chain - m
3	973	27.6	446	A40295	Ig gamma-2a chain
4	952	27.0	220	A31790	Ig kappa chain V r
5	951.5	27.0	469	S37483	Ig gamma-2a chain
6	948	26.9	257	A28664	enterotoxin A prec
7	931	26.4	548	S38864	Ig epsilon chain C
8	929	26.4	246	S38950	Ig gamma chain - m
9	926	26.3	214	S68212	Ig kappa chain (Ma
10	918	26.1	260	C89984	enterotoxin P [imp
11	914.5	26.0	219	S52028	Ig kappa chain - m
12	914	26.0	214	S42402	monoclonal antibod
13	910.5	25.9	219	PC4203	Ig kappa chain (mo
14	903.5	25.7	219	S16112	Ig kappa chain V r
15	900.5	25.6	474	G2MS11	Ig gamma-2b chain
16	896.5	25.5	219	S38865	Ig kappa chain - m
17	892	25.3	218	JC5810	monoclonal antibod
18	890.5	25.3	221	S49220	Ig gamma-1 chain -
19	886.5	25.2	217	S42772	Ig kappa chain - m
20	880	25.0	218	S68241	Ig kappa chain V r
21	878.5	24.9	235	S25058	Ig kappa chain - m
22	876.5	24.9	225	JL0029	Ig kappa chain pre
23	875	24.8	234	S14237	Ig kappa chain pre
24	874	24.8	234	S01320	Ig kappa chain pre
25	873	24.8	475	S01321	Ig gamma-2b chain
26	870	24.7	444	PC4436	monoclonal antibod
27	857	24.3	210	A56169	Ig kappa chain V r
28	853	24.2	240	S06084	Ig kappa chain pre
29	808	22.9	220	S68211	Ig heavy chain (Ma

RESULT 1

A28179 enterotoxin E precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 15-Oct-1999
C:Accession: A28179
R:Couch, J.L.; Soltis, M.T.; Betley, M.J.
J. Bacteriol. 170, 2954-2960, 1988
A:Title: Cloning and nucleotide sequence of the type E staphylococcal enterotoxin gen
A:Reference number: A28179; MUID:88257005; PMID:3384800
A:Accession: A28179
A:Molecule type: DNA
A:Residues: 1-257 <COU>
A:Cross-References: GB:M21319; MID:gl53001; PIDN:AAA26617.1; PID:gl53002
C:Superfamily: enterotoxin B

Query Match 31.4%; Score 1107; DB 2; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-50;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY	226	SKSEBEINKEKLRKSELQGTALGNLKKQIYYNNSKAITSEKSAQDPLNTLLFKGFFTG	285
Db	25	SKSEBEINKEKLRKSELQGNLNLRLQIYYNEKAITENKESDDQFLENTLLFKGFFTG	84
QY	286	HPWYNDLLVDLGSTAAATSEVGSSVDLYGAYGYQCAGGTPNKTACMGVTLHDNNRLT	345
Db	85	HPWYNDLLVDLGSKDATNKYKGVLDLYGAYGYQCAGGTPNKTACMGVTLHDNNRLT	144
QY	346	EKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQ	405
Db	145	EKKVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQ	204
QY	406	RGLIVFHSSEGSTSVSYDLFDAAGQYPTDLLRIYRDNTTISSTLSLSLYLYTT	458
Db	205	RGLIVFHSSEGSTSVSYDLFDAAGQYPTDLLRIYRDNTTISSTLSLSLYLYTT	257

RESULT 2

S37484 Ig kappa chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 11-Jan-2000
C:Accession: S37484
R:Ducancel, F.F.D.
submitted to the EMBL Data Library, February 1993
A:Reference number: S37483
A:Accession: S37484
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-225 <DUC>
A:Cross-References: EMBL:X70424; NID:g406254; PIDN:CAA49869.1; PID:g406255
C:Superfamily: immunoglobulin V region; immunoglobulin homology

RESULT 5

S37483
Ig gamma-2a chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S37483
R:Ducancel, F.F.D.

submitted to the EMBL Data Library, February 1993

A:Reference number: S37483
A:Accession: S37483
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-469 <DUC>
A:Cross-references: EMBL:X70423; NID:g406252; PIDN:CAA9868.1; PID:g406253
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:276-345/Domain: immunoglobulin homology <IMM>

Query Match 27.0%; Score 951.5; DB 2; Length 469;
Best Local Similarity 34.2%; Pred. No. 3.6e-42;
Matches 236; Conservative 62; Mismatches 126; Indels 267; Gaps 13;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYSTGYMHVVKOSPGKGLGWIGRINPNNGVTLY 60

DB 20 QLOQSGGPELVKPGASVKISCKASGYTFDYINNVKPGQGLKGIWIGIPASGNTKY 79

QY 61 NOKFKDKATLTVDKSSSTAYMELRLSITSEDSAVYYCARSTMITNYYMDYWGQTSVTSS 120

DB 80 NENFKCKATLTVDTSSTAYMQLSLITSEDTAVYFCARAMGATATLLDYGQGTTLTVSS 139

QY 121 AKTTPPSVYPLAPGSAQNSMVTGLCLVKGYFPEPVTVTVNSGSLSSGVHFPFAVLQSD 180

DB 140 AKTAPSVYPLAPVCGDGTGSSVTLGCLVKGYFPEPVTVTVNSGSLSSGVHFPFAVLQSD 199

QY 181 LYTLSSTVYVPSSTWPSSEVTCNVAPASSTKVDKIKVPRDSDGSPSEKSEENKDLRKK 240

DB 200 LYTLSSTVYVPSSTWPSQSITCNVAPASSTKVDKIKIEPR-----GPIKPCP 247

QY 241 SELQGTALGNLKOIYYYNKAITSEKSDAOFLTNTLLFKGFTGHPWYNLDLLVLGSTA 300

DB 248 ----- 247

QY 301 ATSEYEGSSVDLYGAYGYQCAGTGNKACMYGGVTLHDNNRLTEKKVPINLWDGKQ 360

DB 248 -----PKRCAPN-----LLGGPSVF----- 263

QY 361 TTVPIDKVKTSKEVTVQELDLQARHYLHGKGLYNSDSFGGKVQRLIVFHSSEGSTVS 420

DB 264 -----IPIPKIKDVLMI-----SLSPVIT 282

QY 421 YDLFDAQGYQPTLLRIYRDNTTISSTLSISLYLTTSIVMTQTPTSLLSAGDRVTIT 480

DB 283 CVVVDVSEDDPD-----VQISWVFNVEVHTAQTOT----- 313

QY 481 CKASQSVNDVAVYQKQKQSPKLLISYTSRYAGVPDRFSGSGYGTDTLTLSISVQAE 540

DB 314 -----HREDYNSTLRVY----- 325

QY 541 AAVYFCQDYNSPPTFGGKTLEIKRAD-----AAPTYSIFPPSSEQLT 584

DB 326 SALPIQHQQWMSGKEP-----KCKVNNKDLPIAPIERTISKPGSVRAPQVYVLPPEEEMT 381

QY 585 SGGASYVCFLENNFYPKDINVKWKIDGSEQRN-----GVLSNWTDDQSDKSTYSMSSTLT 639

DB 382 KQVTLTCVMTDFMPEDIYVEVNTNNGKLTNLYKNTPEVLDS-----DGSYFMYSKLRV 434

QY 640 TKDEYERHNSYCEATHK-TSTSPIVKSPNR 669

DB 435 EKKNNVERNYSYSCVSVHEGLHNHHTTKFSR 465

RESULT 6

A28664

enterotoxin A precursor - Staphylococcus aureus

C:Species: Staphylococcus aureus
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 15-Oct-1999
C:Accession: A28664; A29566
J:Betley, M.J.; Mekalanos, J.J.
J: Bacteriol. 170, 34-41, 1988.
A:Title: Nucleotide sequence of the type A staphylococcal enterotoxin gene.
A:Reference number: A28664; MUID:88086892; PMID:3335483
A:Accession: A28664
A:Molecule type: DNA
A:Residues: 1-257 <BET>
A:Cross-references: GB:M18970; NID:g153120; PIDN:AAA26681.1; PID:g153121
A:Experimental source: strain FRI337
R:Huang, I.Y.; Hughes, J.L.; Bergdoll, M.S.; Schantz, E.J.
J: Biol. Chem. 262, 7006-7013, 1987

A:Title: Complete amino acid sequence of staphylococcal enterotoxin A.
A:Reference number: A29566; MUID:87222293; PMID:3584106
A:Accession: A29566
A:Molecule type: protein
A:Residues: 25-241, S', 243-257 <HUA>
C:Genetics:
A:Gene: entA
A:Map position: 6
C:Superfamily: enterotoxin B

Query Match 26.9%; Score 948; DB 2; Length 257;
Best Local Similarity 76.4%; Pred. No. 2.6e-42;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKSELOQTALGNLKOIYYYNKAITSEKSDAOFLTNTLLFKGFTG 285

DB 25 SEKSEINEKDLRKSELOQTALGNLKOIYYYNKAKTENKESHQDQLQHTILFKGFTD 84

QY 286 HPWYNLDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGTGNKACMYGGVTLHDNNRLT 345

DB 85 HSWYNLDLLVDFDSKDIVDKYKGGKVDLYGAYGYQCAGTGNKACMYGGVTLHDNNRLT 144

QY 346 EEKVPINLWIDGKQTPIDKVKTSKEVTVQELDLQARHYLHGKGLYNSDSFGKVKQ 405

DB 145 EEKVPINLWIDGKQTPLETVKTNKKNVTVQELDLQARRYLQEKYLNLYNSDVPDGKQV 204

QY 406 RGLIVFHSSEGSTVSVDLFDQAQGYQPTLLRIYRDNTTISSTLSISLYLT 458

DB 205 RGLIVFHTSTEPSVNYDLFGAQQYSNTLLRIYRDNKTINSENMHIDIYLYTS 257

RESULT 7

S38864

Ig epsilon chain C region - mouse (fragment)

C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-May-2001
C:Accession: S38864
R:Kipp, B.; Becker, W.; Schlaak, M.
submitted to the EMBL Data Library, November 1993
A:Description: Combination of a defined specificity and desired isotype by cloning of
A:Reference number: S38864
A:Accession: S38864
A:Status: preliminary
A:Molecule type: mRNA

A:Residues: 1-548 <KIP>
A:Cross-references: EMBL:227397; NID:g416537; PIDN:CAA81788.1; PID:g940782
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:353-421/Domain: immunoglobulin homology <IMM>

Query Match 26.4%; Score 931; DB 2; Length 548;
Best Local Similarity 35.0%; Pred. No. 5e-41;
Matches 241; Conservative 83; Mismatches 186; Indels 178; Gaps 19;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYSTGYMHVVKOSPGKGLGWIGRINPNNGVTLY 60

DB 1 QVKLESGGDLVKPGSLKLSAASGLTFSSYGMKSVRQIPDKRLEWATISSGGTYTY 60

QY 61 NOKFKDKATLTVDKSSSTAYMELRLSITSEDSAVYYCARSTMITNYYMDYWGQTSVTSS 120

C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:14-93/Domain: immunoglobulin homology <IMM>

Query Match 25.2%; Score 886.5; DB 2; Length 217;
Best Local Similarity 79.6%; Pred. No. 3.2e-39;
Matches 172; Conservative 14; Mismatches 25; Indels 5; Gaps 1;
QY 461 VMTQPTSLVSGADRVTTICKASQV-----SNDVAVYQKPGQSPKLLISYTSRYAG 515
Db 1 VMTQPSLPLVSLGDQASISCRSSQSLVHTNGNTYLHWLQKPGQSPKLLIYKVSFRSG 60
QY 516 VPDFSGSGYGDTFTLTSSVQAEAAVFCQDYNPPTFGGGTKLEIKRAADAAPTYSI 575
Db 61 VPDFSGSGSGDFTFKISRVEADLGVVFCQSQTVPPTFGSGTKLEIKRAADAAPTYSI 120
QY 576 FPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSRONGVLSWTDQDSDSTYSMS 635
Db 121 FPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSRONGVLSWTDQDSDSTYSMS 180
QY 636 TLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRE 671
Db 181 TLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRE 216

RESULT 20

S68241
Ig kappa chain V region (Mab13-1) - mouse (fragment)
N:Alternate names: immunoglobulin light chain
C:Species: Mus musculus (house mouse)
C:Date: 24-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C:Accession: S68241; S68214
R:Tagagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.
submitted to the EMBL Data Library, March 1994
A:Description: Specific peroxidase activity by formation of an antibody L-chain-porphyrin
A:Reference number: S68241
A:Accession: S68241
A:Molecule type: mRNA
A:Residues: 1-218 <TAK>
A:Cross-references: EMBL:D29670; NID:6473962; PIDN:BAA06141.1; PID:6473963
R:Tagagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.
FEBS Lett. 375, 273-276, 1995
A:Title: Thermostable peroxidase activity with a recombinant antibody L chain-porphyrin
A:Reference number: S68211; MUID:96085223; PMID:7498516
A:Accession: S68214
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 'NI', 3-212 <TAW>
A:Cross-references: EMBL:D29670
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin

Query Match 25.0%; Score 880; DB 2; Length 218;
Best Local Similarity 78.7%; Pred. No. 6.9e-39;
Matches 170; Conservative 13; Mismatches 29; Indels 4; Gaps 1;
QY 460 IVMTQPTSLVSGADRVTTICKASQVSN-----VAVYQKPGQSPKLLISYTSRYAG 515
Db 2 LVLTQSPASLVLGQRATISCRASVSGVYTHWYQKPGQSPKLLISLATNLESG 61
QY 516 VPDFSGSGYGDTFTLTSSVQAEAAVFCQDYNPPTFGGGTKLEIKRAADAAPTYSI 575
Db 62 VPARFSGSGSGDFTLNIHPVEEDVATYQCQSRELPLTFGAGTKLEIKRAADAAPTYSI 121
QY 576 FPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSRONGVLSWTDQDSDSTYSMS 635
Db 122 FPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSRONGVLSWTDQDSDSTYSMS 181
QY 636 TLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRE 671
Db 182 TLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRE 217

RESULT 21

S25058
Ig kappa chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 21-Jan-2000
C:Accession: S25058
R:Fischer, R.; Voss, A.; Niersbach, M.; Munziker, W.; Hirsch, H.J.; Kreuzaler, F.
submitted to the EMBL Data Library, July 1992
A:Description: Production of a Tobacco mosaic virus (TMV) inactivating neotop specific
A:Reference number: S25057
A:Accession: S25058
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-235 <FIS>
A:Cross-references: EMBL:X67211; NID:654828; PIDN:CAA47650.1; PID:654829
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:38-111/Domain: immunoglobulin homology <IMM>
Query Match 24.9%; Score 878.5; DB 2; Length 235;
Best Local Similarity 76.7%; Pred. No. 9e-39;
Matches 171; Conservative 19; Mismatches 32; Indels 1; Gaps 1;

QY 449 LSISLYLTTSIVMTQPTSLVSGADRVTTICKASQVSNVDVAVYQKPGQSPKLLISY 508
Db 13 ISASVILSRGQIVLTQSPAIMSAPGEKVTMTCSASSVSK-MQWYQKSGTSPKRWIYD 71
QY 509 TSSRYAGVPDRFSGSGYGDTFTLTSSVQAEAAVFCQDYNPPTFGGGTKLEIKRAD 568
Db 72 TSKLASGVPRFSGSGSGYSLSLTSSMEADAATYCOOWSSNPLTFGAGTKLEIKRAD 131
QY 569 AAPTYSIFPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSRONGVLSWTDQDSDK 628
Db 132 AAPTYSIFPPSSQLTSGASVVCFLNNYPKIDINVKWIDGSRONGVLSWTDQDSDK 191
QY 629 STYSSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRE 671
Db 192 STYSSSTLTLTKDEYERHNSYTCEATHKTSTSPIVKSFNRE 234

RESULT 22

JL0029
Ig kappa chain precursor (RP93) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 11-Jan-2000
C:Accession: JL0029
J:Chien, N.C.; Pollock, R.R.; Desaymard, C.; Scharff, M.D.
J. Exp. Med. 167, 954-973, 1988
A:Title: Point mutations cause the somatic diversification of IgM and IgG2a antiphosph
A:Reference number: JL0029; MUID:88171315; PMID:3127529
A:Accession: JL0029
A:Molecule type: mRNA
A:Residues: 1-225 <CHI>
A:Experimental source: strain BALB/c, cell line RP93 hybridoma cell
A:Note: the authors translated the codon CGG for residue 106 as Pro, ACC for residue
A:Note: the nucleotide sequence shown is inconsistent with authors' translation becau
ect except for four positions shown above
C:Comment: The protein is an anti-phosphorylcholine antibody.
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: immunoglobulin
F:7-225/Product: Ig kappa chain #status predicted <ILC>
F:7-106/Domain: V region #status predicted <VAR>
F:107-119/Domain: J region #status predicted <JIR>
F:120-225/Domain: C region #status predicted <COR>

Query Match 24.9%; Score 876.5; DB 2; Length 225;
Best Local Similarity 76.4%; Pred. No. 1.1e-38;
Matches 168; Conservative 21; Mismatches 26; Indels 5; Gaps 1;

QY 457 TTSIVMTQPTSLVSGADRVTTICKASQVSN-----VAVYQKPGQSPKLLISYTS 511
Db 5 SSDVLMTQIPLSLPVSLGDQASISCRSSQNIHSTGNTYLEWYLOKPGQSPNLLIYKISN 64
QY 512 RYAGVPDRFSGSGYGDTFTLTSSVQAEAAVFCQDYNPPTFGGGTKLEIKRAADAP 571


```
Qy 415 EGSTVSYDLFDAQGYPTLLRIYRDNNTTISTSLISLSLYLTYTTSIVMTQTPTSLLSVAG 474
Db 299 -----PD-----VQISWFNVNVEVLTAQTQT----- 319
Qy 475 DRVTITCKASQSVSNDVAWYQKPGSKLLISYSSRYAGVPDRFSGSGYGTDFTLTIS 534
Db 320 -----HREDYNSTR 329
Qy 535 SVQA-----EAAVYFCQDYNPPTFGGKLEIKRADAAPTVIFPPSSQQLTSG 586
Db 330 VVSALPIQHDMWSKKEFKCKVKNKNDLPAPITERTISKIGIVRAPQVILPPLPPQLSRK 389
Qy 587 GASVYVCLNNYPKDINVKWKIDGSEQN-----GVLSWTDQDSDKSTYSMSSTLTITLK 641
Db 390 DVSLTCLAVGSPEDISVEWTSNGHTEENYKDTAPVLD-----DGSYFIYSKLNMTK 442
Qy 642 DEYERHNSYTCEATHK 657
Db 443 SKWERTDSFSCNVRHE 458

RESULT 26
PC4436
monoclonal antibody 13-1 heavy chain - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 04-Feb-1998 #sequence_revision 13-Mar-1998 #text_change 21-Jan-2000
C:Accession: PC4436
R:Akashi, S.; Kato, K.; Torizawa, T.; Dohmae, N.; Yamaguchi, H.; Kanachi, M.; Harada, A.
Biochem. Biophys. Res. Commun. 240, 566-572, 1997
A:Title: Structural characterization of mouse monoclonal antibody 13-1 against a porphy-
F:22/Disulfide bonds: JC5810; MUID:98063277; PMID:9398605
A:Accession: PC4436
A:Molecule type: protein
A:Residues: 1-444 <AKA>
C:Comment: This catalytic antibody has peroxidase oxidase activity. It is directed again
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:251-320/Domain: immunoglobulin homology <IMM>
F:22/Disulfide bonds: interchain (to 98) #status predicted
F:99/Disulfide bonds: interchain (to 109) #status predicted

Query Match 24.7%; Score 870; DB 2; Length 444;
Best Local Similarity 32.0%; Pred. No. 5.5e-38;
Matches 214; Conservative 71; Mismatches 131; Indels 252; Gaps 15;
Qy 1 EVLOQSGPDLVKPGASVKISKAGSYGFTGYMHVWKQSPGKGLEWIG--RINPNNGVT 58
Db 1 EVQXVETGGGLVLPNGSLKSLTSGFTFSNYRMHWRQPPGKRLEWATVIVKSDNYGA 60
Qy 59 LYNQKFKDKATLVDSSTAYMELRSITSEDSAYVYCARSTMITNYVNDYWGQGTSTV 118
Db 61 KYAESVGRFTISRDSSKSSVYLQMNRLREEDTATYYCCRTPWV--YAMDCWGQGTSTV 118
Qy 119 SSKATTPPSVPLAPGSAQTNSWYTLGCLVKGYFPEPVTVWNSGSLSSGHHVTPAVLQ 178
Db 119 SSKATTPPSVPLAPGSAQTNSWYTLGCLVKGYFPEPVTVWNSGSLSSGHHVTPAVLQ 178
Qy 179 SLDYTLSSSVTPSPSETVTCNVAHPASSTKYVDKIKVPRDSDGPGSEKSEINEKDLR 238
Db 179 SLDYTLSSSVTPSPSETVTCNVAHPASSTKYVDKIKVPRDCG----- 223
Qy 239 KKSELOGTALGNLKIYYNNSKAITSSEKSDQFNTLLFKGFTGHPWYNLDLVDLGS 298
Db 224 --CKPICITVPEVSSVFIFPPK-----PKDVLITL----- 252
Qy 299 TAATSEYEGSSVDLGGAYGYGACAGTGNKTKACMYGGVTLHDNNRLTBEKKVPINLWDG 358
Db 253 -----TP-KVFCVVVDIS-----KDDPEVQFSWFVD- 277
Qy 359 KOTTVPIDKVTSSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVGGLI 409
Db 278 ---DVEVHTAQTPREEQFNSTFRSVSELPIMHODWLNKKEFKCRVNSAAPPAPTEK--- 331
Qy 410 VPHSSEGSTVSYDLFDAQGYPTLLRIYRDNNTTISTSLISLSLYLTYTTSIVMTQTPTSL 469
```

```
Db 332 -----TISKT----- 336
Qy 470 LVSAGDRVTITCKASQSVSNDVAWYQKPGSKLLISYSSRYAGVPDRFSGSGYGTDF 529
Db 337 ----- 336
Qy 530 TLTISVQAEDAAYVFCQDYNPPTFGGKLEIKRADAAPTVIFPPSSQQLTSGGAS 589
Db 337 -----KGRPKAPOVYTIPTPKQMAKDKVS 361
Qy 590 VVCFNNYPKDINVKWKIDGSEQNGLVNSWTDQDSDKSTYSMSSTLTITLKDEYERHNS 649
Db 362 LTCMITDFPEDITVEWQWNGOPAEV-NKNTQPIMDT-DGSYFVYSKLVNQKSNWEAGT 419
Qy 650 YTCEATHK 657
Db 420 FTCSVLHE 427

RESULT 27
A56169
Ig kappa chain V region (clone 23.2) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 11-Jan-2000
C:Accession: A56169
R:Monfardini, C.; Kieber-Emmons, T.; Vontfeldt, J.M.; O'Malley, B.; Rosenbaum, H.; God
J. Biol. Chem. 270, 6628-6638, 1995
A:Title: Recombinant antibodies in bioactive peptide design.
A:Reference number: A56169; MUID:95204454; PMID:7896802
A:Accession: A56169
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-210 <MON>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 24.3%; Score 857; DB 2; Length 210;
Best Local Similarity 79.4%; Pred. No. 1e-37;
Matches 166; Conservative 13; Mismatches 26; Indels 4; Gaps 1;
Qy 450 IVMTQPTSLVSGADRVITTCASQSQSNDSND---VAVYQKPGSKPLLIYSYSSRYAG 515
Db 2 IVLTQSPATLSVSLQRATISCRASKSVSSSGYSYMHVYQKPGQPKVLIYLAQNSLESG 61
Qy 516 VPDRESGSGYGTDFTLTISVQAEDAAYVFCQDYNPPTFGGKLEIKRADAAPTVSI 575
Db 62 VPPRESGSGSGTDFTLNHPVEEADAATYYCOHSRELPTWTEGGGTREIKRADAAPTVSI 121
Qy 576 FPPSSEQLTSGGASVVCFLNNYPKDINVKWKIDGSEQNGLVNSWTDQDSDKSTYSMS 635
Db 122 FPPSSEQLTSGGASVVCFLNNYPKDINVKWKIDGSEQNGLVNSWTDQDSDKSTYSMS 181
Qy 636 TLTITLKDEYERHNSYTCEATHKSTSPIV 664
Db 182 TLTITLKDEYERHNSYTCEATHKSTSPIV 210

RESULT 28
S06084
Ig kappa chain precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 21-Jan-2000
C:Accession: S06084
R:Crower, J.S.; Smith, M.A.; Cooper, H.J.
Nucleic Acids Res. 17, 7992, 1989
A:Title: Nucleotide sequence of Y3-Ag 1.2.3. rat myeloma immunoglobulin kappa chain c
A:Reference number: S06084; MUID:90016888; PMID:2508067
A:Accession: S06084
A:Molecule type: mRNA
A:Residues: 1-240 <CRO>
A:Cross-references: EMBL:X16129; NID:956457; PIDN:CAA34256.1; PID:956458
C:Superfamily: immunoglobulin V region; immunoglobulin homology
```

C:Keywords: heterotetramer; immunoglobulin
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-240/Product: Ig kappa chain #status predicted <MAT>
F:153-222/Domain: immunoglobulin homology <IMM>

Query Match 24.2%; Score 853; DB 2; Length 240;
Best Local Similarity 69.6%; Pred. No. 1.9e-37;
Matches 165; Conservative 31; Mismatches 37; Indels 10; Gaps 2;

QY 445 STSLSISLVLYTT---SIVMTQPTSLVLSAGDRVTITCKASQSV-----SNDVAVY 494
Db 3 SQTQVLSLLILISGTCGDFVMTQSPSSLAVASGETVITNCKSSQLFSGNOKNTLAWY 62
QY 495 QOKPQSPKLLISYTSRYAGVDRFSGSGYGTDFTLTSSVQAEDAAVYFCQDDNSPP 554
Db 63 QOKPGOSPCLLIYWASTROSGVDRFIGSGSGTDFTLTSSVQAEDLAIVYCLQYYETPY 122
QY 555 TFGGGTGLKIKRADAPTYSIFPPSEQLTSGGASVVCFLNFYKIDINVKKIDGSRQ 614
Db 123 TFGAGTKLEKLRADAPTYSIFPPSTEQLATGGASVVCFLMNFYPRDISVKWKIDGTERR 182
QY 615 NGVLNSWTQDDSKDSTYSMSSTLTITKDEYERHNSYTCEATHTKSTPIVKSFNRNE 671
Db 183 DGVLDSTVDQDSDKSTYSMSSTLSLSKADYESHNLITCEVHKTSPPVKSFNRNE 239

RESULT 29
Ig heavy chain (Mab13-1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 29-Jul-1997 #sequence_revision 29-Aug-1997 #text_change 20-Jun-2000
C:Accession: S68211
R: Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.
FBS Lett. 375, 273-276, 1995
A:Title: Thermostable peroxidase activity with a recombinant antibody L chain-porphyrin
A:Reference number: S68211; MUID:96085223; PMID:7498516
A:Status: preliminary; nucleic acid sequence not shown
A:Accession: S68211
A:Molecule type: mRNA
A:Residues: 1-220 <TAK>
A:Cross-references: EMBL:D29669; NID:g473958; PIDN:BAA06140.1; PID:g473959
A:Note: the sequence of residues 1-4, 213-220 and the corresponding nucleotide sequence
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:136-200/Domain: immunoglobulin homology <IMM>

Query Match 22.9%; Score 808; DB 2; Length 220;
Best Local Similarity 70.2%; Pred. No. 3.7e-35;
Matches 153; Conservative 24; Mismatches 37; Indels 4; Gaps 2;

QY 6 QSGPDLVPGASVKISCKASGYSTGYMHVWKQSPGKLEWIG--RINPNNGVTLYNOK 63
Db 2 ESGGGLVRPGNSLKLISLTSGTFSNYRMHLRQPPGKRLKLEAVITVKSVDNKGAKYAES 61
QY 64 FKDKATLTVDKSTTAYMELRSLTSEDSAVYVCARSTMTITNVMDYWGOGTSTVYSSAKT 123
Db 62 VGRFTITRSDSKSVYLQNMNLRREEDATYCCRTPMV--YAMDCWGOGTSTVYSSAKT 119
QY 124 TTPSVYPLAPGSAATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFTFPAVLQSDLYT 183
Db 120 TTPSVYPLAPGSAATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFTFPAVLQSDLYT 179
QY 184 LSSSVTVPSSTWSPSETVTCNVAHPASSTKVKDKKIIPRD 221
Db 180 LSSSVTVPSSTWSPSETVTCNVAHPASSTKVKDKKIIPRD 217

RESULT 30
S68213
Ig heavy chain (Mab03-1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 29-Jul-1997 #sequence_revision 17-Sep-1997 #text_change 21-Jan-2000
C:Accession: S68213
R: Takagi, M.; Kohda, K.; Hamuro, T.; Harada, A.; Yamaguchi, H.; Kamachi, M.; Imanaka, T.

FBS Lett. 375, 273-276, 1995
A:Title: Thermostable peroxidase activity with a recombinant antibody L chain-porphyrin
A:Reference number: S68211; MUID:96085223; PMID:7498516
A:Accession: S68213
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-213 <TAK>
A:Cross-references: EMBL:D29667
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:137-201/Domain: immunoglobulin homology <IMM>

Query Match 20.9%; Score 737.5; DB 2; Length 213;
Best Local Similarity 64.7%; Pred. No. 1.6e-31;
Matches 139; Conservative 29; Mismatches 44; Indels 3; Gaps 1;

QY 2 VOLQOQSPDLVPGASVKISCKASGYSTGYMHVWKQSPGKLEWIGRINPNNGVTLYN 61
Db 2 VOLVESGGGLVPGGSRKLSCAASGFTFSFGFHHVWRQAPKEGKLEWVAYISSGSSSIYXA 61
QY 62 QKFKDKATLTVDKSTTAYMELRSLTSEDSAVYVCARSTMTITNVMDYWGOGTSTVYSSA 121
Db 62 DTVKGRETISRDNPKNTLFLQMTSLASEDTAMYCARSWLLP--FDYWGOGTTLTVSSA 118
QY 122 KTTTPSVYPLAPGSAATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFTFPAVLQSD 181
Db 119 KTTTPSVYPLAPGCGDITGSSVTGLCLVKGYFPEPVTVTWNSGSLSSSVHFTFPALLQSG 178
QY 182 YTLSSSVTVPSSTWSPSETVTCNVAHPASSTKVKDKK 216
Db 179 YTMSSSVTVPSSTWSPSETVTCNVAHPASSTTVDDK 213

RESULT 31
PC4155
Ig gamma-2b chain V-C region MabB23 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 21-Jan-2000
C:Accession: PC4155
R: Kwak, J.W.; Choi, B.K.; Lee, D.I.; Kang, Y.K.; Seo, Y.G.; Cho, W.K.; Han, M.H.
Gene 169, 237-239, 1996
A:Title: Cloning and characterization of cDNAs coding for heavy and light chains of a
A:Reference number: PC4155; MUID:96194809; PMID:8647454
A:Accession: PC4155
A:Molecule type: mRNA
A:Residues: 1-231 <KWA>
A:Cross-references: GB:U28970; NID:g1262180; PIDN:AAC52489.1; PID:g1262181
A:Note: This protein has unusual amino acid compared with the conserved sequences of
C:Superfamily: immunoglobulin C region; immunoglobulin homology
F:1-231/Product: heavy chain #status predicted <MAT>
F:98-102/Region: unique D sequence
F:103-119/Region: V region
F:139-203/Domain: immunoglobulin homology <IMM>

Query Match 20.3%; Score 715; DB 2; Length 231;
Best Local Similarity 60.9%; Pred. No. 2.5e-30;
Matches 137; Conservative 29; Mismatches 55; Indels 4; Gaps 2;

QY 1 EVQLQOQSPDLVPGASVKISCKASGYSTGYMHVWKQSPGKLEWIGRINPNNGVTLY 60
Db 1 EVQLVESGGLVAPQSLSITCTVSGFSLTDYGVSWIRQPPGKLEWLIWA-GGSTFY 59
QY 61 NQKFKDKATLTVDKSTTAYMELRSLTSEDSAVYVCARSTMTITNVMDYWGOGTSTVYSS 120
Db 60 NSALKSRSLINKDNKSKSQVFLKMNLSLHTDDTAMYCVKIHEDRYDWFVWAGTITTVYSS 119
QY 121 AKTTPSVYPLAPGSAATNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHFTFPAVLQSD 180
Db 120 AKTTPSVYPLAPRCGDTTGSSTVTLGCLVKGYFPEPVTVTWNSGSLSSSVHFTFPALLQSG 179
QY 181 LYTLSSTVTVPSSTWSPSETVTCNVAHPASSTKVKDKKIIPRDSGGP 225
Db 180 LYTMSSTVTVPSSTWSPSETVTCNVAHPASSTTVDDKLEP---SGP 221

```
RESULT 32
JE0244
Ig kappa chain NIG2 precursor - human
C:Species: Homo sapiens (man)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: JE0244
R:Alim, M.A.; Hara, Y.; Hossain, M.S.; Takeda, K.; Yamagata, F.; Yamaki, S.; Kazi, H.; et al.
submitted to JIPID, November 1998
A:Description: A new subgroup of k type light chains (VKV) identified in cases of AL amyloidosis
A:Reference number: JE0243
A:Accession: JE0244
A:Molecule type: protein
A:Residues: 1-215 <ALI>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:16-90/Domain: immunoglobulin homology <IMM>

Query Match 19.9%; Score 702.5; DB 2; Length 215;
Best Local Similarity 62.1%; Pred. No. 1e-29;
Matches 133; Conservative 36; Mismatches 42; Indels 3; Gaps 2;

QY 460 IYVMTPTSLVSGADRVITTCASQSVNDVAVYQKQPGQSPKLLISYTSRYAGVDPDR 519
Db 2 VLVTSPTSLVSPGERATLSCRASQSVHNLAWYQKQPGQAPRLLIYRSTRATGIPAR 61
QY 520 FSGSGYGTDFLTITISSVQAEADAAYFCQDYNS--PPTFGGQTKLEIKRADAAPTVSIFP 577
Db 62 FSGSGGTDFLTITSSLOEEDFALYCCQ-YNTWPLTFGGGQTKVEIKRTVAAPSVFIFP 120
QY 578 PSEQLTSGASVCFVCLNNFYPKDINVKWKIDGSRQNGVLNSWTDQDSKDYTSMSSTL 637
Db 121 PSDEQLKSGTAVVCLLNNFYPREAKVQKVDNALQSGNSQESVTEQDSKDYTSLSLSTL 180
QY 638 TLTKDEYERHNSYTCETHKTSPTIVKSFNRNE 671
Db 181 TLTKADYEAHKYACEVTHQGLSSPVTKSFNRGE 214

RESULT 33
S04845
Ig heavy chain precursor - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 26-Aug-1999
C:Accession: S04845; S05695
R:Amemiya, C.T.; Haire, R.N.; Litman, G.W.
Nucleic Acids Res. 17, 5388, 1989
A:Title: Nucleotide sequence of a cDNA encoding a third distinct xenopus immunoglobulin
A:Reference number: S04845; MUID:89345103; PMID:2503814
A:Accession: S04845
A:Molecule type: mRNA
A:Residues: 1-549 <AME>
A:Cross-references: EMBL:X15114
R:Litman, G.W.
submitted to the EMBL Data Library, April 1989
A:Reference number: S05695
A:Accession: S05695
A:Molecule type: mRNA
A:Residues: 'LC', 3-308, 'H', 310-549, <LIT>
A:Cross-references: EMBL:X15114; NID:g64799; PID:g763031
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: glycoprotein; heterotetramer; immunoglobulin
F:26-109/Domain: immunoglobulin homology <IMM>
F:281,294/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 19.9%; Score 700.5; DB 2; Length 549;
Best Local Similarity 28.0%; Pred. No. 4.2e-29;
Matches 194; Conservative 107; Mismatches 214; Indels 179; Gaps 23;

QY 1 EYVLOQSGPDVLVKPGASVKISCKASGYPTGYMHVWVKQSPGKLEIWRINPNNGVTLY 60
Db 12 DIELVQPSSEIKSPGESIKLSCKTSCYFTNTWIIHWIQVPGKGLQWIGRIYPGDADTDY 71
QY 61 NOKFKDKATLTVDKSTTAYMELRSLTSDSAVYYCARSTMTITNYVMDYWGQGTSTVSS 120
```

```
Db 72 SSSYQGRCHISTDNPQSTTFQLNNLKVEDTAIYYCAREGV--GVVFDYWGQGTWVTVTS 129
QY 121 AKTTPPSVYPLAP--GSAQTNSMWTLCCLVKGYPEPVTWNSGSLSSGVHTTPAVL- 177
Db 130 AFLHAPSVFPLPCPCGSSS--SDSHVTIGCLFPLAPVDVKNWNSGTSITSGLNKFPVLIQ 188
QY 178 QSDLTSLSSVTPVSTWPS--ETVTCNVAHPASSTKVDKKIVPRDSSGGGSEKSEINEKD 236
Db 189 QSGLFASSQLTIPISDKWAKKSFECNVEHKPTSKYVKQKIECQDEPEPIETVEI---- 244
QY 237 LRKKELOQTALGNLKIYYIYNSKAITSEKSADQFLTNLTLLFKGFTTGHWPYNOLLVDL 296
Db 245 -----LQG----- 247
QY 297 GSTAATSEYEGSSVDLYGAYGYCAGGTPNKTACMYGGVTLHDNNRLTEERKVPINLWI 356
Db 248 -----PCASSKSVELLCLITGYA-----PSEIKVHWLL 275
QY 357 DGKQTTVPIDKVKTSKE-----VTVOELDLQARHYLHGKFLGYS--DSFGGKVQR 406
Db 276 NQOVNTI SPSNKKPCKEENGTFSSRSKVSVPKED-----WNSEDSYCKVYTH 322
QY 407 GLIVFHS--SEGTVSYDLFDAOQYPTDLLRIYRONTTISSTLSISLYLYTTSI--VMTQ 464
Db 323 P--ASHTKTEASTKKC-----DETAI-----TPKVDVLP 350
QY 465 TPTSLVSGADRVITTCASQSVNDVAVYQKQPGQSPKLLISYTSRYAGVDPDRFSGSG 524
Db 351 SPKDLVLT--KEAKYVCVISRMASTDDLTQVQWSRDKKAL-----AFDSAPEKAYDGT 402
QY 525 YCTDFTLTITSSVQAEADAAYFCQ---QDYNSPPTFGGQTKLEIKRAD---AAPTVSIFP 578
Db 403 FTVKSTLISPGDWENKKQFNCKVVPDLPS-----IEKSIQSQDPCTETITLLP 456
QY 579 SSEQLTSGASVCFVCLNNFYPKDINVKWKIDGSRQNGVLNSWTD--QDSKDYTSMSSTL 637
Db 457 SDELRNDFISLICMLKNFRPQDIYVFNKQGVLTLEEDYVMTTTPVLEEEEGFISFSL 516
QY 638 TLTKDEYERHNSYTCETHKTSPTIVKSFNRNE 671
Db 517 TIARSDWMRGATYSYCIAAHNTISQRDIRK--NRGK 549

RESULT 34
JE0242
Ig kappa chain NIG26 precursor - human
C:Species: Homo sapiens (man)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: JE0242
R:Alim, M.A.; Yamaki, S.; Hossain, M.S.; Takeda, K.; Kojima, M.; Takashi, I.; Shinoda
submitted to JIPID, November 1998
A:Description: Structure relationship of kappa type light chains with AL amyloidosis:
A:Reference number: JE0241
A:Accession: JE0242
A:Molecule type: protein
A:Residues: 1-215 <ALI>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:16-91/Domain: immunoglobulin homology <IMM>

Query Match 19.7%; Score 695.5; DB 2; Length 215;
Best Local Similarity 63.1%; Pred. No. 2.3e-29;
Matches 135; Conservative 32; Mismatches 44; Indels 3; Gaps 3;

QY 460 IYVMTPTSLVSGADRVITTCASQSVND--VAVYQKQPGQSPKLLISYTSRYAGVDP 518
Db 2 IYVTSQPTSLVSPGERATLSCRASQSVNNYLAAYQKQPGQAPSLIYDASSRATGIPD 61
QY 519 RFGSGSGYGTDFLTITSSVQAEADAAYFCQDYNSPP--TFGGGQTKLEIKRADAAPTVSIFP 577
Db 62 RFGSGSGGTDFLTITSGLEPDAFYICQO--YDRPWTTFGGGQTKVEIKRTVAAPSVFIFP 120
QY 578 PSEQLTSGASVCFVCLNNFYPKDINVKWKIDGSRQNGVLNSWTDQDSKDYTSMSSTL 637
```

Db 121 PSDEQLKSTASVCLNNFYPREAKVQKVDNALQSGNSQESVTEQDSKDYSTLSSTL 180
Qy 638 TLTKEDEYERHNSYTCEATHKTKSTSPIVKSFNRNE 671
Db 181 TLSKADYEKKHYACEVTHQGLSSPVTKSFNRGE 214
RESULT 35
B31790
Ig heavy chain V region (17/9) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 23-May-1997
C:Accession: B31790
R:Schulze-Gahmen, U.; Rini, J.M.; Arevalo, J.; Stura, E.A.; Kenten, J.H.; Wilson, I.A.
J. Biol. Chem. 263, 17100-17105, 1988
A:Title: Preliminary crystallographic data, primary sequence, and binding data for an an
A:Reference number: A92686; MUID:89034213; PMID:3182835
A:Accession: B31790
A:Molecule type: mRNA
A:Residues: 1-254 <SCH>
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:15-98/Domain: immunoglobulin homology <IMM>
Query Match 19.4%; Score 683.5; DB 2; Length 254;
Best Local Similarity 63.8%; Pred. No. 1.2e-28;
Matches 146; Conservative 22; Mismatches 48; Indels 13; Gaps 8;
Qy 1 EVQLQSGDPLVKPKASKVSKSGYFTGYMHVWVQSPKGLGWIRPNNGVTLY 60
Db 1 EVQLVSGDPLVKPGSKLSCAASGFSSYGMVSRQTPDKRLIEWATISNGGGYTY 60
Qy 61 NQKFKDKATLVTDKSTAYMELRLSLTSDSAVYVCARSTMITNVMYVMDVWGOGTSTVYSS 120
Db 61 PDSVKGRTISRDNKNTLYLQMSLKSDSAMYCARRERYDENGFAVWGOGTLVTVSA 120
Qy 121 AKTTPSVVPLAP--GSAAQTNMVTGLCLVKGYFPEPVTV-TW---NSGSLSSG-VHT 172
Db 121 AKTTPSVVPLAPVCGXXDTTGSSTGLCLVKGYFPEPVTLTWXXXNSGSLSSGXXVHT 180
Qy 173 FPAVLQS--DLVTLSSSVTPVSS-TWP-SETVT-CNVAHPASSTKVDKK 216
Db 181 FPAVLQXXDLVTLSSSVTPVSS-TWPSQSITFCXNVAHPASSTKVDKK 229
RESULT 36
JE0243
Ig kappa chain NIG93 precursor - human
C:Species: Homo sapiens (man)
C:Date: 05-Dec-1998 #sequence_revision 05-Dec-1998 #text_change 21-Jan-2000
C:Accession: JE0243
R:Allm, M.A.; Hara, Y.; Hossain, M.S.; Takeda, K.; Yamagata, F.; Yamaki, S.; Kazi, H.; T
submitted to JIPID, November 1998
A:Description: A new subgroup of k type light chains (VkV) identified in cases of AL amy
A:Reference number: JE0243
A:Accession: JE0243
A:Molecule type: protein
A:Residues: 1-215 <ALI>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:16-90/Domain: immunoglobulin homology <IMM>
Query Match 19.3%; Score 679.5; DB 2; Length 215;
Best Local Similarity 61.0%; Pred. No. 1.6e-28;
Matches 130; Conservative 33; Mismatches 49; Indels 1; Gaps 1;
Qy 460 IVMTQPTSLVYAGDRVTITCKASQSVSNDAVYQKPGQSPKLLISYTSRYAGVPDR 519
Db 2 IVMTQSPATLSVSPGERATLSCRASQSVATNVVMYMQKLGQAPRLLIYDASTRATGVPAR 61
Qy 520 FSGSGYGTDTLTITSSVQAEDAAYVFCQDYNSPTFGGTRKLEIKRADAAPTVSIFPPS 578
Db 62 FSGSGGTFTLTITSSLOSEFAIYYCQHNNAWPPTFGGTRKRTVAAAPSVFIAPP 121

Qy 579 SSEQLTSGASVYVCFLLNFFPKDINVKWKIDGERQNGVLNSWTQDSKDYSTLSSTLT 638
Db 122 SDEQLKSGTASVVCLLNFFYPREAKVQKVDNALQSGNSQESVTEQDSKDYSTLSSTLT 181
Qy 639 LTKDEYERHNSYTCEATHKTKSTSPIVKSFNRNE 671
Db 182 LSKADYEKKHYACEVTHQGLSSPVTKSFNRGE 214
RESULT 37
S33161
Ig kappa chain - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 20-Jun-2000
C:Accession: S33161
R:Foley, R.C.; Beh, K.J.
submitted to the EMBL Data Library, July 1990
A:Description: Isolation and characterisation of sheep kappa light chain cDNA.
A:Reference number: S33161
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-230 <FOL>
A:Cross-references: EMBL:X54110; NID:g297103; PIDN:CAA38046.1; PID:g1364221
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:143-212/Domain: immunoglobulin homology <IMM>
Query Match 19.2%; Score 675; DB 2; Length 230;
Best Local Similarity 59.4%; Pred. No. 2.9e-28;
Matches 126; Conservative 37; Mismatches 49; Indels 0; Gaps 0;
Qy 460 IVMTQPTSLVYAGDRVTITCKASQSVSNDAVYQKPGQSPKLLISYTSRYAGVPDR 519
Db 18 IQVTSPLSLASLTERVSIITCQSQSVSNLYNWYQKPGQAPKLLIYATRLHTDVPDR 77
Qy 520 FSGSGYGTDTLTITSSVQAEDAAYVFCQDYNSPTFGGTRKLEIKRADAAPTVSIFPPS 579
Db 78 FSGSGGTDTLTITSNLEANDATYYCLQYESTPLAFGGGTNVEIKRDAQSVFLFKPS 137
Qy 580 SQQLTSGASVYVCFLLNFFPKDINVKWKIDGERQNGVLNSWTQDSKDYSTLSSTLT 639
Db 138 EQQLTGTVTSVVCLVNDVFKDINVKVGVGTQNSFNQNSFTQDSKDYSTLSSTLT 197
Qy 640 TKDEYERHNSYTCEATHKTKSTSPIVKSFNRNE 671
Db 198 SSSEIQSHNAYACEVSHKSLPTALVKSFNKNE 229
RESULT 38
S22080
Ig heavy chain precursor (B/Mt.4A.17.H5.A5) - bovine
N:Alternate names: Ig gamma-1 chain C region (clone 8.10)
C:Species: Bos primigenius taurus (cattle)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999
C:Accession: S22080; S06610; A31303
R:Sanders, P.G.
submitted to the EMBL Data Library, November 1991
A:Reference number: S22080
A:Accession: S22080
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-470 <SAS>
A:Cross-references: EMBL:X62916; NID:g439; PIDN:CAA44699.1; PID:g440
R:Symons, D.B.A.; Clarkson, C.A.; Beale, D.
Mol. Immunol. 26, 841-850, 1989
A:Title: Structure of bovine immunoglobulin constant region heavy chain gamma 1 and g
A:Reference number: S06610; MUID:90097956; PMID:2513487
A:Accession: S06610
A:Molecule type: DNA
A:Residues: 142-470 <SYM>
A:Cross-references: EMBL:X16701

C>Date: 03-Mar-1994 #sequence_revision 03-Mar-1994 #text_change 16-Jul-1999
C:Accession: A49444
R:Saul, F.A.; Poljak, R.J.
Proteins 14, 363-371, 1992
A:Title: Crystal structure of human immunoglobulin fragment Fab New refined at 2.0 A resolution
A:Reference number: A49444; MUID:93066153; PMID:1438175
A:Accession: A49444
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-220 <SAU>
A>Note: sequence modified after extraction from NCBI backbone
A>Note: this sequence report includes corrections based on crystal structure refinement
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin
F:137-202/Domain: immunoglobulin homology <IMM>

Query Match 17.3%; Score 611; DB 2; Length 220;
Best Local Similarity 54.3%; Pred. No. 5.6e-25;
Matches 120; Conservative 36; Mismatches 61; Indels 4; Gaps 3;

QY 1 EVLOQSGPDLVKPGASVKISCKASCSFTGYMHVWVKSPGKGLWIGRIINPNNGVTLY 60
DB 1 QVLESGGGLVLRPSQTLSTCVSTSFDDIYTWVWPPGKLEWIGIVF-YTGTTLL 59

QY 61 NQKFKDKATLVDKSSTAYMELRLSTSDSAVYICARSTMITNYMYDYGQTSVTSS 120
DB 60 DPSLRGRVTMLVNTSKNQFSRLSSVTAADTAVYICARNLIAGG--IDVWGQSLVTSS 117

QY 121 AKTTPSVYPLAPGSAQNTNSMTGLCKLVKGFPEPVTWNSGSLSSGSHVFPVAVLQSD 180
DB 118 ASTKGSVYPLAPSSKSTSGGTAAALGCKLVKGFPEPVTWNSGALTSVGTFFPAVLQSS 177

QY 181 -LYLTSSVTPSPSTWPTVCNVAHPASSTKVDKVIKVP 220
DB 178 GLYSLSSVTPSSSLGTQYICNVNHHKPSNTKYDKRKEPK 218

RESULT 45
A33953
enterotoxin D precursor - Staphylococcus aureus
C:Species: Staphylococcus aureus
C>Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 15-Oct-1999
C:Accession: A33953
R:Bayles, K.W.; Iandolo, J.J.
J. Bacteriol. 171, 4799-4806, 1989
A:Title: Genetic and molecular analyses of the gene encoding staphylococcal enterotoxin
A:Reference number: A33953; MUID:89359112; PMID:2549000
A:Accession: A33953
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-258 <BAV>
A:Cross-references: GB:M28521; NID:g1492109; PIDN:AAB06195.1; PID:g758691
C:Superfamily: enterotoxin B

Query Match 17.3%; Score 611; DB 2; Length 258;
Best Local Similarity 51.1%; Pred. No. 6.9e-25;
Matches 118; Conservative 35; Mismatches 78; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLKKSELOGTALGNLQKIYYNSKAITSEKSADQFLTNLLFGKFFTG 285
DB 26 NEMIDSVKELHKKSELSSTALNNMKHSYADKNPIIGENKSTGDFLENTLLYKKFFTD 85

QY 286 HPWYNLLVDLSTAATSEVGSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNRILT 345
DB 86 LINFEDLLINFNSKEMAQHFKSKNVDDYPIRVSINCYGGEIDRTACTYGGVTPHEGNKIL 145

QY 346 EEKKVPINLWDGKQTPVPIDKVKTSKKEVTVQELQARHYLHGKFLYNSDSFGKVKQ 405
DB 146 ERKKIPINLWINGVQKESLDKVTQDKNVTVQELDAQARRYLQDKLYLNNDTLGGKIQ 205

QY 406 RGLVPHSSEGSTVSDLFDAQGPDPDLLRIYRDNTTSSLSLSLYL 456
DB 206 RGKIEFSDSGSKSVYDLEFDVKGDFPEKQLRIYSNKTLSLTHELHIDYLY 256

RESULT 46
S14683
Ig mu chain precursor, membrane-bound (clone 201) - human
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 23-Jul-1999
C:Accession: S14683; S08047
R:Friedlander, R.M.; Nussenzweig, M.C.; Leder, P.
Nucleic Acids Res. 18, 4278, 1990
A:Title: Complete nucleotide sequence of the membrane form of the human IgM heavy chain
A:Reference number: S14683; MUID:90332450; PMID:2115996
A:Accession: S14683
A:Molecule type: mRNA
A:Residues: 1-627 <FRI>
A:Cross-references: EMBL:X17115; NID:g33450; PIDN:CAA34971.1; PID:g33451
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: immunoglobulin; membrane protein
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-627/Product: Ig mu chain #status predicted <MAT>
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 16.9%; Score 595.5; DB 2; Length 627;
Best Local Similarity 26.8%; Pred. No. 1.3e-23;
Matches 190; Conservative 100; Mismatches 209; Indels 211; Gaps 27;

QY 1 EVLOQSGPDLVKPGASVKISCKASCSFTGYMHVWVKSPGKGLWIGRIINPNNGVTLY 60
DB 20 QVOLVOSGAEVKRPSSVKSCAGGTSSYAISWVRQAPGOGLEWGGIIPFGTANY 79

QY 61 NQKFKDKATLVDKSSTAYMELRLSTSDSAVYICARSTMITNYV-----M 107
DB 80 AQKFGQRTITADESTSTAYMELSLRSEDYAVYICAKTGILGYPSSGHWPNSDYYVYGM 139

QY 108 DYWGQGTSTVTSVSAKTPPSVYPL--APGSAQAQNTSMWTLGCLVKGFPEPVTW---N 162
DB 140 DVWGQGTSTVTSVSGSASAPTLFPLVSCENSPSDTSS--VAVGCLAQDFLPSITFSKYN 198

QY 163 SGLSSGVHTFPVAVLQSDLYLTSSSVTPSS---TWPSETVTCNVAHPASS----- 210
DB 199 NSDISS--TRGFPVSLRGKGYAATQVLLPSKDVMOGTDEHVYCKVOHPNGNKEKNVPLPV 257

QY 211 -----TKVDKIKVPRDS--GGPSEKSENEKDLRKKSELOGTALGNLQKIYYNSKAIT 263
DB 258 IAEPLPKVSFVPPRDPFGNPRSKSLICQATGSPRQIOWSLREGQV----- 308

QY 264 SSEKSADQFLTNLLFGKFFTHPWYNDLLDGLGTAATSEYEGSSVDLYGAYGYQCAG 323
DB 309 -----GSGVTTDQVQAEAKE-----SG 325

QY 324 GTPNKTACMYGGVTLHDNRNLTEKKVPIINLWDGKQTPVPIDKVKTSKKEVTVQELDLQ 383
DB 326 PTYKVT---STLTIKESD-----WLSOSMFTCRVD----- 353

QY 384 ARHYLHGKFLYNSDSFGKVGQRLIVFHSSEGSTVSYDLFDAQGOYPTLLRIYRDNTT 443
DB 354 -----HRGL-TFOQNASSMCVPD-----QDTARV-----A 379

QY 444 ISSTLSLSLYLYTTSIVMTQPTSTLLYSAGDRVTITCKASQSVNSDVAVYQOKPGQSPK 503
DB 380 IPPSFASIFL---TKSTKLTLCLVTDL--TTYDSVTI-----SWTRQN-GEAVK 421

QY 504 LLISYTSRRYAGVPDRFSGSGYGTDTLLTSSVQAEADAAYVFCQODYNSPPTF----- 556
DB 422 ---THTNISEHPNATFSAVG-----EASI---CEDWNSEGRFTCTVTHT 461

QY 557 --GGGTKLEIKRAAAA---PTVSIFPPSPSEQLT--SGGASVVCFLNNFYPKIDINVKWKID 609
DB 462 DLPSPLKQTIISRPKGVALLHRPDPVYLLPPAREQLNLRSATITCLVTGTFSPADVFVQWMQR 521

QY 610 GS--ERQNGVLNSWTDQDSKSTYSMSSTPLTLTKDEYERHNSYTCEATHK 657
DB 522 GQPLSPEKYVTSAPWPEQAPGRYFAHSILTVSEEWNTGETYTCVVAHE 571

F:143-212/Domain: immunoglobulin homology <IMM>

```

RESULT 47
PT0219
Ig kappa chain V-C region (PLC18) - pig (fragment)
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 11-Jan-2000
C:Accession: PT0219
R:Lammers, B.M.; Beaman, K.D.; Kim, Y.B.
Mol. Immunol. 28, 877-880, 1991
A>Title: Sequence analysis of porcine immunoglobulin light chain cDNAs.
A:Reference number: PT0219; MUID:91342694; PMID:1715030
A:Accession: PT0219
A:Molecule type: mRNA
A:Residues: 1-178 <LAM>
A:Cross-references: GB:M59321; NID:q164508; PIDN:AAA03520.1; PID:g164509
A:Experimental source: spleen, strain Minnesota Miniature
A>Note: The authors translated the codon CTC for residue 141 as Ser
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
F:1-70/Domain: V region (fragment) <VRG>
F:12-18/Region: complementarity-determining 1
F:19-51/Region: framework 1
F:52-60/Region: complementarity-determining 2
F:61-70/Region: framework 2
F:71-178/Domain: C region <CRG>
F:96-156/Disulfide bonds: #status predicted
F:176/Disulfide bonds: interchain #status predicted

Query Match 16.8%; Score 591; DB 2; Length 178;
Best Local Similarity 64.6%; Pred. No. 4.7e-24;
Matches 113; Conservative 23; Mismatches 39; Indels 0; Gaps 0;

QY 497 KPGQSPKLLISYTSRYAGVPPRFSGSGVGDTFTLTISVQAEADAAVFCQDYNSPPNF 556
DB 1 KPGQSPQLLIVASDRASGVPDRFGSGSGDTFTLKINSVEADAGVYCHQKFEPRTF 60

QY 557 GGGTKLEIKRADAAPTVSIFPPPSSEQLTSGGASVVCFLNFPKIDINVKWKIDGSGRQNG 616
DB 61 GGGTKLEIKRADAKSVIFPPPSKSQLATPTVSVCLINFFPREISVKWVDGVVQSSG 120

QY 617 VLNSWTDQSKDSTYSMSSTLTLTDEYERHNSYCEATHKSTSPIVKSFNRNE 671
DB 121 HPDSVTEQDSKSTYSLSLSTLPTSOYLSHNLNLSYCEVTHKTLASPLVTSFNRNE 175

RESULT 48
A20969
Ig kappa chain precursor V-J-C regions - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C>Date: 03-Aug-1990 #sequence_revision 03-Aug-1990 #text_change 21-Jan-2000
C:Accession: A20969; A25448
R:McCartney-Francis, N.; Skurja Jr., R.M.; Mage, R.G.; Bernstein, K.B.
Proc. Natl. Acad. Sci. U.S.A. 81, 1794-1798, 1984
A>Title: Kappa-chain allotypes and isotypes in the rabbit: cDNA sequences of clones encoding expression.
A:Reference number: A20969; MUID:84170388; PMID:6424124
A:Accession: A20969
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-229 <MCC>
A:Cross-references: GB:K01359; NID:q165373; PIDN:AAA31334.1; PID:g165374
R:Akimenko, M.A.; Marilame, B.; Rougeon, F.
Proc. Natl. Acad. Sci. U.S.A. 83, 5180-5183, 1986
A>Title: Evolution of the immunoglobulin kappa light chain locus in the rabbit: evidence
A:Reference number: A94110; MUID:86259753; PMID:3088570
A:Contents: Ig kappa-1 chain, 69 allotype, J-K1.2 segment
A:Accession: A25448
A:Molecule type: DNA
A:Residues: 111-123 <AKI>
A:Cross-references: GB:M14067; GB:M14062; GB:M14063; GB:M14064; GB:M14065; GB:M14066; NT
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
```

```

Query Match 16.2%; Score 570.5; DB 2; Length 229;
Best Local Similarity 53.0%; Pred. No. 7.4e-23;
Matches 114; Conservative 35; Mismatches 63; Indels 3; Gaps 3;

QY 459 SIVMTQTPTSLVSAGDRVTITCKASQSVSNDAVYQQRPGQSPKLLISYTSRVRAGVDP 518
DB 15 ALVMTQTPASVSAVGGTGTIKQASENIYSSLAWYQQRPGQPKLLIYGASTLASGVPS 74

QY 519 RFGSGYGTDFTLTISVQAEADAAVFC-QQDYNSPPTFGGKLEIK-RADAAPTVSTP 576
DB 75 RFKGRSGTEYTLTISGVQREDAATYYCLGSDSSDSTAFCGGTELEILCDPPIAPTPLVF 134

QY 577 PPSSQLTSGGASVVCFLNFPKIDINVKWKIDGSRONGVLNSWTDQSKDSTYSMSST 636
DB 135 PPSAQLTETVTIVCVANKFRPNDITVTWKVDDELQQSGIENSTTPQSPEDCTYNLSST 194

QY 637 LTLTKDEYERHNSYCEATHKSTSPIVKSFNRNE 671
DB 195 LSLTKAQYNSHVSYYTCEVH-NSGSAIVQSFNRGD 228

RESULT 49
S29594
Ig gamma chain (MM65) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 05-Nov-1999
C:Accession: S29594
R:Seymour, R.
submitted to the EMBL Data Library, February 1991
A:Reference number: S29593
A:Accession: S29594
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-178 <SEY>
A:Cross-references: EMBL:X57857; NID:g52590; PIDN:CAA40992.1; PID:g52591
C:Keywords: immunoglobulin

Query Match 16.0%; Score 562; DB 2; Length 178;
Best Local Similarity 73.5%; Pred. No. 1.5e-22;
Matches 108; Conservative 13; Mismatches 24; Indels 2; Gaps 1;

QY 1 EVQLQQSGDPLVKPGASVKISCKASGYFTGYMHVWKQSPGKGLFWIGRINPNNGVTLY 60
DB 14 EVQLQQSGAEVLKPGASVKLSCTASGFNIKDTYIHHVKQRPQKGLWIGRIDPANGYTEY 73

QY 61 NQKFKDKATLTVDKSTTTAYMELRLSLTSDSAVYYCARSTMITNYYMDYWGQGSTVTVSS 120
DB 74 DPKFGKATITADTSTNTAYLQLSLTSEDTAIVYCTGNYA--YGM DYWGQGSTVTVSS 131

QY 121 AKTTPPSVYPLAPGSAQAQNSMVTLCG 147
DB 132 AKTTPSVYPLAPGSAQAQNSMVTLCG 158

RESULT 50
I54782
gene Pvt-1a/Ig-Ck protein - mouse (fragment)
C:Species: Mus sp. (mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 02-Sep-2000
C:Accession: I54782
R:Huppi, K.; Siwarski, D.
Int. J. Cancer 59, 848-851, 1994
A>Title: Chimeric transcripts with an open reading frame are generated as a result of
A:Reference number: I54782; MUID:95080867; PMID:7989128
A:Accession: I54782
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-126 <RES>
A:Cross-references: GB:S76258; NID:g913277; PIDN:AAB32752.1; PID:g913278
C:Genetics:
A:Gene: Pvt-1a/Ig-Ck
```

C:Superfamily: pre-B cell omega light chain; immunoglobulin homology

Query Match 15.9%; Score 561; DB 2; Length 126;
Best Local Similarity 93.0%; Pred. No. 1.le-22;
Matches 107; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 557 GGGTKLEIKRADAPTIVSIFPPSSQLTSGGASVVCFLNNFYPKDINVKWKIDGSRQNG 616
Db 11 GAAAKVKSIIWADAAPTIVSIFPPSSQLTSGGASVVCFLNNFYPKDINVKWKIDGSRQNG 70
QY 617 VLNSWTDQDSKDSYMSSTLTLTDEYERHNSYTCETHKSTSPIVKSFNRNE 671
Db 71 VLNSWTDQDSKDSYMSSTLTLTDEYERHNSYTCETHKSTSPIVKSFNRNE 125

Search completed: January 8, 2003, 11:58:57
Job time : 32.2464 secs


```

FT METAL 211 211 ZINC (BY SIMILARITY).
FT METAL 249 249 ZINC (BY SIMILARITY).
FT METAL 251 251 ZINC (BY SIMILARITY).
SQ SEQUENCE 257 AA; 29358 MW; 27EDA94B97770CE3 CRC64;

Query Match 31.4%; Score 1107; DB 1; Length 257;
Best Local Similarity 89.7%; Pred. No. 9.5e-62;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKSELOCTALGNLKOIYYNSKAITSEKSADOFLENTLFLKGFFTG 285
DB 25 SEKSEINEKDLRKSELOCTALGNLKOIYYNEKAITENKESDDQFLENTLFLKGFFTG 84
QY 286 HPWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCACMGVTLHDNNRLT 345
DB 85 HPWYNLLVDLGSKDNTNKKYKKKVDLYGAYGYQCAGGTPNKTCACMGVTLHDNNRLT 144
QY 346 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGKQVQ 405
DB 145 EEKKVPINLWIDGKQTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGKQVQ 204
QY 406 RGLIVFHSSEGSVSYDLDPDAQOQYPTDLLRIYRDNTTISSTSLSLSLYLTT 458
DB 205 RGLIVFHSSEGSVSYDLDPDAQOQYPTDLLRIYRDNTTINSENHLIDLXYTT 257

RESULT 2
ETXA_STAAW STANDARD; PRT; 257 AA.
AC FI1363;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type A precursor (SEA).
GN ENTA OR MW1889.
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12043478;
RA Baba T., Takeuchi F., Kuroda H., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=PRI337;
RX MEDLINE=88086892; PubMed=3335483;
RA Bettley M.J., Mekalanos J.J.;
RT "Nucleotide sequence of the type A staphylococcal enterotoxin gene.";
RL J. Bacteriol. 170:34-41(1988).
RN [3]
RP SEQUENCE OF 25-257.
RX MEDLINE=87222293; PubMed=3584106;
RA Huang I.-Y., Hughes J.L., Bergdoll M.S., Schantz E.J.;
RT "Complete amino acid sequence of staphylococcal enterotoxin A.";
RL J. Biol. Chem. 262:7006-7013(1987).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=95354648; PubMed=7628431;
RA Schad E.M., Zaitseva I., Zaitsev V.N., Dohlsten M., Kalland T.,
RA Schlievert P.M., Ohlendorf D.H., Svensson L.A.;
RT "Crystal structure of the superantigen staphylococcal enterotoxin
RT type A.";
RL EMBO J. 14:3292-3301(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=97113025; PubMed=8943278;

```

```

RA Sundstroem M., Hallen D., Svensson A., Schad E., Dohlsten M.,
RA Abrahmsen L.;
RT "The Co-crystal structure of staphylococcal enterotoxin type A with
RT Zn2+ at 2.7-A resolution. Implications for major histocompatibility
RT complex class II binding.";
RL J. Biol. Chem. 271:32212-32216(1996).
RN [6]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=96022987; PubMed=7552730;
RA Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
RT "Residues defining V beta specificity in staphylococcal
RT enterotoxins.";
RL Nat. Struct. Biol. 2:680-686(1995).
RN [7]
RP COMPARISON OF STRUCTURE OF SEA AND SEC2.
RX MEDLINE=97334373; PubMed=9191070;
RA Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;
RT "A structural and functional comparison of staphylococcal
RT enterotoxins A and C2 reveals remarkable similarity and
RT dissimilarity.";
RL J. Mol. Biol. 269:270-280(1997).
CC -!- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
CC for the toxin interaction with MHC class II.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY A BACTERIOPHAGE.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP004828; BAB95754.1;
DR EMBL: M18970; AAX26681.1;
DR PIR: A28664; A28664.
DR PIR: A29566; A29566.
DR PDB: 1ESF; 11-JUL-96.
DR PDB: 1SEA; 15-OCT-95.
DR PDB: 1SXT; 19-NOV-97.
DR InterPro: IPR001961; Stap/Strep_toxin.
DR Pfam: PF01123; Stap_Strp_toxin; 1.
DR Pfam: PF02876; Stap_Strp_tox_C; 1.
DR PRINTS: PR00279; BACTRLTOXIN.
DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;
KW 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 257 ENTEROTOXIN TYPE A.
FT DISULFID 120 130
FT METAL 211 211 ZINC.
FT METAL 249 249 ZINC.
FT METAL 251 251 ZINC.
FT CONFLICT 242 242 T -> S (IN REF. 3).
SQ SEQUENCE 257 AA; 29669 MW; ADEBF5BCA1F14677 CRC64;

Query Match 26.9%; Score 948; DB 1; Length 257;
Best Local Similarity 76.4%; Pred. No. 6.4e-52;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKSELOCTALGNLKOIYYNSKAITSEKSADOFLENTLFLKGFFTG 285
DB 25 SEKSEINEKDLRKSELOCTALGNLKOIYYNEKAITENKESDDQFLENTLFLKGFFTG 84
QY 286 HPWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCACMGVTLHDNNRLT 345
DB 85 HPWYNLLVDLGSKDNTNKKYKKKVDLYGAYGYQCAGGTPNKTCACMGVTLHDNNRLT 144

```

[illegible][illegible]

RESULT 4			
GCI_MOUSE	STANDARD;	PRT;	324 AA.
ID	GCI_MOUSE		
AC	P01868;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	30-MAY-2000 (Rel. 39, Last annotation update)		
DE	Ig gamma-1 chain C region.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Muridae; Mus		

OC Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.
 NCBI_TaxID=10090;
 [1]
 RN SEQUENCE FROM N. A.
 RP MEDLINE=80045036; PubMed=115593;
 RX Honjo T., Obata M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
 RA Takahashi N., Mano Y.;
 RT "Cloning and complete nucleotide sequence of mouse immunoglobulin
 gamma 1 chain gene";
 RL Cell 18:559-568(1979).

[2]
RN
RP
RX
RX
RA
RA
RA

0314 167955 500(12177).
SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
MEDLINE=80202559; PubMed=6769752;
Obata M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
Yamawaki K., Gotoh S., Ohtsuka H., Nakamura T., Ohno S.

RA Mano Y., Seidman J.G., Peterlin B.M., Leder P., Honjo T.:
 RT "Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
 RT cloned in a bacterial plasmid.";
 RT Gene 9:87-97 (1980).
 RL
 RN [3]
 RN
 RP SEQUENCE OF 70-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
 RX MEDLINE=80012837; PubMed=113776;
 RX Rogers J., Clarke P., Saisser W.:
 RT "Sequence analysis of cloned cDNA encoding part of an immunoglobulin

RT nucleic acid analysis of cloned cDNA encoding part of an immunoglobulin heavy chain";

RL Nucleic Acids Res. 6:3305-3321(1979).

RA [4]

RP SEQUENCE (MYELOMA PROTEIN MOPC 21).

RX MEDLINE=78242288; PubMed=98524;

RA Aдетубо К.:

RT "Evolution of immunoglobulin subclasses. Primary structure of a murine myeloma gammal chain".

RL
RN
RP
RS
RX
RY
RA
RB
RC
RD
RE
RF
RG
RH
RI
RJ
RK
RL
RM
RN
RO
RP
RQ
RR
RS
RT
RU
RV
RW
RX
RY
RZ

J. Biol. Chem. 253:6088-6075 (1978).
[5]
DISULFIDE BONDS (MOPC 21).
MEDLINE=73008889; PubMed=5073237;
RA Svasti J., Milstein C.;
RT "The disulfide bridges of a mouse immunoglobulin G1 protein.";
RI Blochem. J. 126:837-850(1972).
CC -----
CC This SWISS-PROT entry is copyright It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC

Query Match	Score 611;	Score 17.3%;	Length 258;	DB 1;	Length 258;


```
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00407; IgC1; 1.
DR PROSITE: PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 26 86
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
SQ SEQUENCE 106 AA; 11778 MW; 4B51FF5EF49BAEB5 CRC64;

Query Match 15.8%; Score 556; DB 1; Length 106;
Best Local Similarity 100.0%; Pred. No. 3.4e-28;
Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 567 ADAAPTSTLTKDEYRHNSTCEATHTKSTSPVKSFRNE 671
DB 1 ADAAPTSTLTKDEYRHNSTCEATHTKSTSPVKSFRNE 60
QY 627 KDSTYSMSSTLTKDEYRHNSTCEATHTKSTSPVKSFRNE 671
DB 61 KDSTYSMSSTLTKDEYRHNSTCEATHTKSTSPVKSFRNE 105

RESULT 7
GCL_RAT
ID GCL_RAT STANDARD; PRT; 326 AA.
AC P20759;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-1 chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR; PS0017; PS0017.
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam; PF00047; Ig_3.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 112 HINGE.
FT DOMAIN 113 219 CH2.
FT DOMAIN 220 326 CH3.
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 306
FT CARBOHYD 176 176 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 326 AA; 35946 MW; 013BAB45EF49B9DA CRC64;

Query Match 15.6%; Score 550; DB 1; Length 326;
Best Local Similarity 27.7%; Pred. No. 3.6e-27;
Matches 150; Conservative 64; Mismatches 90; Indels 238; Gaps 14;

QY 121 AKTTSPSYVPLAPGSAQTNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHPTFPAVLQSD 180
DB 1 AETAPSVYPLAPGALKNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHPTFPAVLQSG 60
```

```
QY 181 LYTSLSSVTPSPSTWPSSETVTCNVAHPASSTKVDKKIIVPRDSGGPSEKSEINEKDLRKK 240
DB 61 LYTSLSSVTPSPSTWPSSETVTCNVAHPASSTKVDKKIIVPRDSGGPSEKSEINEKDLRKK 109
QY 241 SELQGTALGNLKOIYYNSKAITSSSEKSDAQFLNTLTLFKGFFTGHPWVNDLLDLGSTA 300
DB 110 -----ICGTSEVSS-----VFIFPKPKDVL----- 130
QY 301 ATSEYEGSSVDLYGAYGYOCAGGTPNKACMYGGVTLHDNNRLTEKKVPIINLWIDG-- 358
DB 131 -----TITL-----TP-KVTCVVVDIS-----QDDPEVHFSWFVDDVE 162
QY 359 ---KOTTVPIDKVTSSKEVTVQELDLQARHYLHGKFLGFLYNSDSFGGKVQORGLIVFHSSE 415
DB 163 VHTAOTRPEEPQFNSTFR--SVSELPILHQLDNLGR----- 196
QY 416 GSTVSYDLFDAOGQYPTDLLRIYRDNTTISSTLSISLYLTTSIVMTQTPTSLVSAGD 475
DB 197 ----- 196
QY 476 RVTIITCKASQSYSDVNAVYQKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDFLTITSS 535
DB 197 ---TFRCKV-----TSAAFPSPIEKTISK 217
QY 536 VOAEADAAYVFCOODYNSPPTFGGTTKLEIKRADAAPTVSIFPPSSEOLTSGASVVCFLN 595
DB 218 PE-----GRTOV-----PHVYTMSPTEEMTQNEVSIITCMVK 249
QY 596 NPYPKDINVKWKIDGERQVLSNWTDOOSKDSYSSMSSTLTCLKDEYRHNSTCEAT 655
DB 250 GFYPDIIYVQMNGQPQEN-YKNTPPTMDT-DGSVFLYSLKNVKKKQWQGNQTFCTSVL 307
QY 656 HK 657
DB 308 HE 309

RESULT 8
GCA_RAT
ID GCA_RAT STANDARD; PRT; 322 AA.
AC P20760;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2A chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M13804; AAA41376.1; ALT_INIT.
DR PIR; PS0019; PS0019.
DR HSSP; P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_c1.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00407; IgC1; 2.
DR PROSITE; PS00290; Ig_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
```

```
FT NON_TER 1 1
FT DISULFID 27 82
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 136 196
FT DISULFID 242 300
FT CARBOHYD 172 172
SQ SEQUENCE 322 AA; 35186 MW; E8EA136A9DE01EDB CRC64;

Query Match 15.3%; Score 539; DB 1; Length 322;
Best Local Similarity 28.2%; Pred. No. 1.7e-26;
Matches 153; Conservative 55; Mismatches 92; Indels 242; Gaps 15;

QY 121 AKTTPPSVPLAPGSAQTNSMTLGLCVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
Db 1 AETTAPSVPLAPGTALKNSMTLGLCVKGYFPEPVTVTWNSGALSSGVHTFPAVLQSG 60

QY 181 LYTLSSTVPSSTWPSSETVTCNVAHPASSTKVDKKIVPRDSGGPSEKSEINEKDLRKK 240
Db 61 LYTLSSTVPSSTWSSQAQVTCNVAHPASSTKVDKKIVPREC----- 102

QY 241 SELQGTALGNLQIYYNKAITSSEKSADQFLTNLTKGFFTGHPWYNDLLVLDGSTA 300
Db 103 -----NPCGCTGSEVS-----VFIPPTKDKVL----- 126

QY 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKACMYGGVTLHDNNRLTEKKVPIINLWIDG-- 358
Db 127 -----TITL-----TP-KVTCVVVDISQND-----PEVRESWIDDOVE 158

QY 359 ---KOTTVPIDKVKTSKEVTVQELDLQARHLYHCKGLGKLYNSDSFGKVGRLVHFHSSE 415
Db 159 VHTAQTHAPEKOSNTLR--SVSELPVHROWLNGK-----TFKCKVN----- 199

QY 416 GSTVSVDLFDAGQVDPDLLRIYRDNNTISSTLSLSLYLTYSIVMTQTPSLVLSAGD 475
Db 200 -----SGAPPAI----- 207

QY 476 RVTITCKASQSVNDVAWYQKPGSKPLISYTSRYAGVDPDRFSGSGYGTDFLTISS 535
Db 208 -----EKSIS-----KPEGTPTP----- 219

QY 536 VQAEADAAVYFCQDYNSPPTFGGKTKLEIKRADAAPTVSIIPPSSSEQITSGASVVCFLN 595
Db 220 -----GPOVYTMAPPKEEMTQSQVSTICAVK 245

QY 596 NFYPKIDINVKKIDGSRQNGVLNWTQDQSKDSYSMSSTLTLLTKDEYERNHSYTCBAT 655
Db 246 GFYPDPDIYTEKMKNGQPQEN-YKNTPTPMDT-DGSYFLYSLKLVNKKETWQQNGTFTCSVL 303

QY 656 HK 657
Db 304 HE 305

RESULT 9
ID GCAA_MOUSE STANDARD; PRT; 330 AA.
AC P01863;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-2A chain C region, A allele.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=81076554; PubMed=6777755;
RA Sikorav J.-L., Auffray C., Rougeon F.;
RT "Structure of the constant and 3' untranslated regions of the murine
Ba1b/c gamma 2a heavy chain messenger RNA.";
```

```
RL Nucleic Acids Res. 8:3143-3155(1980).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=81198976; PubMed=6262729;
RA Yanawaki-Katsuka Y., Miyata T., Honjo T.;
RT "The complete nucleotide sequence of mouse immunoglobulin gamma 2a gene
and evolution of heavy chain genes: further evidence for intervening
sequence-mediated domain transfer.";
RL Nucleic Acids Res. 9:1365-1381(1981).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=81223894; PubMed=6787604;
RA Olio R., Auffray C., Morchamps C., Rougeon F.;
RT "Comparison of mouse immunoglobulin gamma 2a and gamma 2b chain genes
suggests that exons can be exchanged between genes in a multigenic
family.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:2442-2446(1981).
RN [4]
RP MYELOMA PROTEIN MOPC 173.
RX MEDLINE=74175517; PubMed=4831970;
RA Bourgois A., Fougereau M., Rocca-Serra J.;
RT "Determination of the primary structure of a mouse IgG2a
immunoglobulin:amino-acid sequence of the Fc fragment. Implications
for the evolution of immunoglobulin structure and function.";
RL Eur. J. Biochem. 43:423-435(1974).
RN [5]
RP DISULFIDE BONDS.
RX MEDLINE=73056887; PubMed=4565406;
RA de Preval C., Fougereau M.;
RT "Determination of the primary structure of a mouse gamma G2a
immunoglobulin. Identification of the disulfide bridges.";
RL Eur. J. Biochem. 30:452-462(1972).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V00798; CAA24178.1; -.
DR PIR; A02152; G2MSA.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; IG_like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
KW NON_TER 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT MOD_RES 330 330 REMOVED POST-TRANSLATIONALLY.
SQ SEQUENCE 330 AA; 36389 MW; B84361C5445A6864 CRC64;

Query Match 14.0%; Score 491.5; DB 1; Length 330;
Best Local Similarity 26.4%; Pred. No. 1.5e-23;
Matches 151; Conservative 43; Mismatches 110; Indels 267; Gaps 13;

QY 121 AKTTPPSVPLAPGSAQTNSMTLGLCVKGYFPEPVTVTWNSGSLSSGVHTFPAVLQSD 180
Db 1 AKTTAPSVPLAPVCGDITGSSVTLGLCVKGYFPEPVTLTWNSGSLSSGVHTFPAVLQSD 60

QY 181 LYTLSSTVPSSTWPSSETVTCNVAHPASSTKVDKKIVPRDSGGPSEKSEINEKDLRKK 240
Db 1 LYTLSSTVPSSTWPSSTWSSQAQVTCNVAHPASSTKVDKKIVPREC----- 102
```

```
Db 61 LYTSSSVTVTSWPSSQITCNVAHPASSTKVKKIEPR-----GPTIKPCP----- 108
QY 241 SELOGTALGNLQKIYYNSKAITSSSEKSDQFLTNTLLFKGFTGHPWYNLLVDLGSTA 300
Db 109 ----- 108
QY 301 ATSEYEGSVLDYGAYGYQCAGGTPNKTACMYGVTLHDNNRLTEKKVPINLWIDGKQ 360
Db 109 -----PCKCPAPN-----LLGSPSVF----- 124
QY 361 TTVPIDKVKTSKEVTVOELDLOARHLYHGKEGLYNSDSFGCKVQORGLIVFHSSEGSTVS 420
Db 125 -----IFFPKIKDVLMI-----SLSPIVT 143
QY 421 YDLFDAQGYDPTLLRIYRDNTTISSTLSISLYLTTSIVMTQPTPTSLLSVAGDRVTIT 480
Db 144 CVVDVSEDDPD-----VQISWFVNNVEVHTAQQT----- 174
QY 481 CKASOSVNDVAVYQOKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTISVQAE 540
Db 175 -----HREDYNSTLRVV----- 186
QY 541 AAVYFCQDYNSPPTFGGKTLEIKRAD-----AAPTYSIFPPSSEQILT 584
Db 187 SALPTQHQDWMGSKEF-----CKVNNKDLPAPIERTISKPGSVRAPQVYVLPPEEEMT 242
QY 585 SGGASVVFCLNFFPKDINVKWKIDGSRQN-----GVLSNWTQDQSDKSTYSMSSTLT 639
Db 243 KKOYTLTCMTWDFMPEDIYVENNKGKTELYNKTEPVLDSE-----DGSYFMYSKLVR 295
QY 640 TKDEYERHNSYTCEATHK-TTSPIVKSFNR 669
Db 296 EKKNWERNYSYCSVVHEGLHNHHTTKFSR 326

RESULT 10
GCAM_MOUSE
ID GCAM_MOUSE STANDARD; PRT; 399 AA.
AC P01865;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE I9 gamma-2A chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains."
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS. IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-329 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM OF
CC THE A ALLELE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00471; AAB59661.1; ALT_INIT.
```

```
DR PIR; A02154; G2MSAM.
DR HSSP; P01842; 7FAB.
DR MGD; MG1:96443; Igh-1.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IGC1; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT TRANSMEM 346 363
FT DOMAIN 364 399
FT CARBOHYD 180 180
FT SEQUENCE 399 AA; 44020 MW; 4C38138BFAED3FF0 CRC64;
Query Match 14.0%; Score 491.5; DB 1; Length 399;
Best Local Similarity 26.4%; Pred. No. 2e-23;
Matches 151; Conservative 43; Mismatches 110; Indels 267; Gaps 13;
QY 121 AKTTPSPVPIAPGSAQTNSMTLGLVKGYPPEVTVTNWNSGLSSGVHTFPVAVLQSD 180
Db 1 AKTTAPSVPIAPVCGDGTGSSVTLGLVKGYPPEVTVTNWNSGLSSGVHTFPVAVLQSD 60
QY 181 LYTSSSVTVSPSWSETVTCNVAHPASSTKVKKIIVPROSGGSPSEKSEENKDLRKK 240
Db 61 LYTSSSVTVTSWPSSQITCNVAHPASSTKVKKIEPR-----GPTIKPCP----- 108
QY 241 SELOGTALGNLQKIYYNSKAITSSSEKSDQFLTNTLLFKGFTGHPWYNLLVDLGSTA 300
Db 109 ----- 108
QY 301 ATSEYEGSVLDYGAYGYQCAGGTPNKTACMYGVTLHDNNRLTEKKVPINLWIDGKQ 360
Db 109 -----PCKCPAPN-----LLGSPSVF----- 124
QY 361 TTVPIDKVKTSKEVTVOELDLOARHLYHGKEGLYNSDSFGCKVQORGLIVFHSSEGSTVS 420
Db 125 -----IFFPKIKDVLMI-----SLSPIVT 143
QY 421 YDLFDAQGYDPTLLRIYRDNTTISSTLSISLYLTTSIVMTQPTPTSLLSVAGDRVTIT 480
Db 144 CVVDVSEDDPD-----VQISWFVNNVEVHTAQQT----- 174
QY 481 CKASOSVNDVAVYQOKPGQSPKLLISYTSRYAGVDPDRFSGSGYGTDTFTLTISVQAE 540
Db 175 -----HREDYNSTLRVV----- 186
QY 541 AAVYFCQDYNSPPTFGGKTLEIKRAD-----AAPTYSIFPPSSEQILT 584
Db 187 SALPTQHQDWMGSKEF-----CKVNNKDLPAPIERTISKPGSVRAPQVYVLPPEEEMT 242
QY 585 SGGASVVFCLNFFPKDINVKWKIDGSRQN-----GVLSNWTQDQSDKSTYSMSSTLT 639
Db 243 KKOYTLTCMTWDFMPEDIYVENNKGKTELYNKTEPVLDSE-----DGSYFMYSKLVR 295
QY 640 TKDEYERHNSYTCEATHK-TTSPIVKSFNR 669
Db 296 EKKNWERNYSYCSVVHEGLHNHHTTKFSR 326

RESULT 11
HV12_MOUSE
ID HV12_MOUSE STANDARD; PRT; 117 AA.
```

```
AC P01756;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Ig heavy chain v region MOPC 104E.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE AND CARBOHYDRATE-LINKAGE SITE ASN-55.
RX MEDLINE=83075344; PubMed=6816276;
RA Kehry M.R., Fuhrman J.S., Schilling J.W., Rogers J., Sibley C.H.,
RA Hood L.E.;
RT "Complete amino acid sequence of a mouse mu chain: homology among
RT heavy chain constant region domains.";
RL Biochemistry 21:5415-5424(1982).
CC -!- MISCELLANEOUS: THE SEQUENCE OF THE LIGHT CHAIN OF THIS IGM MYELOMA
CC -!- PROTEIN HAS ALSO BEEN DETERMINED.
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.
DR PIR; A02039; MHMS4E.
DR HSSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin v region; Glycoprotein.
FT DISULFID 22 96 BY SIMILARITY.
FT CARBOHYD 55 55 N-LINKED (GLCNAC. . .) (COMPLEX).
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12983 MW; 3CF8ACE4BE447E41 CRC64;

Query Match 13.7%; Score 481.5; DB 1; Length 117;
Best Local Similarity 78.3%; Pred. No. 1.6e-23;
Matches 94; Conservative 8; Mismatches 15; Indels 3; Gaps 1;

QY 1 EVLOQSGPDLVKPGASVKISCKASYSTGYGMHWKQSPGKLEWIGRINPNNGVTL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 EVLOQSGPELVKPGASVKMSCKASYSTGYGMHWKQSHGKSLWIGINPNNGVTSY 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 NQKFKDKATLVDSKSTTAYMELRSLTSDSAVYICARSTMTINYMDYWGQTSVTSS 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 61 NQKFKDKATLVDSKSTTAYMQLNSLTSDSAVYICARD---YDFYDVGAGTTVTSS 117
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 12
HV07_MOUSE STANDARD; PRT; 139 AA.
AC P01751; P01752;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain v region B1-8/186-2 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=81234348; PubMed=6798376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the Npb family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:623-637(1981).
CC -!- MISCELLANEOUS: THE B1-8 MU CHAIN MRNA WAS CLONED FROM A HYBRIDOMA
CC MAKING ANTIBODIES TO THE HAPTEN (4-HYDROXY-3-NITROPHENYL)ACETYL
CC (NPB ANTIBODIES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
```

```
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00529; AAA38170.1; -
DR PIR; A02034; MHMS18.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin v region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 139 IG HEAVY CHAIN V REGION B1-8/186-2.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DOMAIN 118 124 D SEGMENT.
FT DOMAIN 125 139 JH2 SEGMENT.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 139 139
SQ SEQUENCE 139 AA; 15719 MW; 1B57DD4FDC09F465 CRC64;

Query Match 13.7%; Score 481; DB 1; Length 139;
Best Local Similarity 74.2%; Pred. No. 2.1e-23;
Matches 89; Conservative 14; Mismatches 17; Indels 0; Gaps 0;

QY 1 EVLOQSGPDLVKPGASVKISCKASYSTGYGMHWKQSPGKLEWIGRINPNNGVTL 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 20 QVLOQPGAEVLKPGASVKLSCKASYSTGYGMHWKQSPGKLEWIGRIDPNSGGTKY 79
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 NQKFKDKATLVDSKSTTAYMELRSLTSDSAVYICARSTMTINYMDYWGQTSVTSS 120
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB 80 NQKFKDKATLVDSKSTTAYMQLNSLTSDSAVYICARYDYGGSYFDYWGQGTTLTVSS 139
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 13
HV13_MOUSE STANDARD; PRT; 117 AA.
AC P01757;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain v region J559.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=80078170; PubMed=6765983;
RA Schilling J., Clevinger B., Davie J.M., Hood L.;
RT "Amino acid sequence of homogeneous antibodies to dextran and DNA
RT rearrangements in heavy chain v-region gene segments.";
RL Nature 283:35-40(1980).
CC -!- MISCELLANEOUS: THE SEQUENCES OF 10 HYBRIDOMA PROTEINS THAT ALSO
CC BIND DEXTRAN DIFFER FROM THAT SHOWN AT 1-7 POSITIONS, MANY OF
CC WHICH OCCUR IN THE D AND J SEGMENTS.
CC -!- MISCELLANEOUS: THIS PROTEIN BINDS DEXTRAN.
DR PIR; A26242; MHMSJ5.
DR HSSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin v region.
FT DISULFID 22 96 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 13024 MW; 292E2AF4BE447E41 CRC64;
```

```
Query Match 13.6%; Score 480.5; DB 1; Length 117;
Best Local Similarity 78.3%; Pred. No. 1.8e-23;
Matches 94; Conservative 7; Mismatches 16; Indels 3; Gaps 1;

QY 1 EVOLQSGPDLVPGASVKISKASGYSTFGYYMHVWKSPGKLEWIGRINPNNGVTLY 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 EVOLQSGPDLVPGASVKISKASGYSTFGYYMHVWKSPGKLEWIGRINPNNGVTLY 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 NQKFKKATLVTKSSATTAYMELRLSTSDSAVYYCARSTMTITNYMDYWGQCTSVTVSS 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 NQKFKKATLVTKSSATTAYMELRLSTSDSAVYYCARSTMTITNYMDYWGQCTSVTVSS 117
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 14
HV51_MOUSE STANDARD; PRT; 118 AA.
ID HV51_MOUSE STANDARD; PRT; 118 AA.
AC P06330;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region AC38 205.12.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE.
RX MEDLINE=84182519; PubMed=6201362;
RA Dildrop R., Bovens J., Stekevitz M., Beyreuther K., Rajewsky K.;
RT "A V region determinant (idiotope) expressed at high frequency in B
RT lymphocytes is encoded by a large set of antibody structural genes.";
RL EMBO J. 3:517-523(1984).
DR PIR: A02040; MHMS38.
DR HSSP: P01789; IMCP.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 98 V SEGMENT.
FT DOMAIN 99 104 D SEGMENT.
FT DOMAIN 105 118 J SEGMENT.
FT DISULFID 22 96 BY SIMILARITY.
FT NON_TER 118 118
SQ SEQUENCE 118 AA; 12934 MW; 94F7BEE4C762A018 CRC64;

Query Match 13.6%; Score 480; DB 1; Length 118;
Best Local Similarity 79.2%; Pred. No. 2e-23;
Matches 95; Conservative 6; Mismatches 17; Indels 2; Gaps 1;

QY 1 EVOLQSGPDLVPGASVKISKASGYSTFGYYMHVWKSPGKLEWIGRINPNNGVTLY 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 EVOLQSGPDLVPGASVKISKASGYSTFGYYMHVWKSPGKLEWIGRINPNNGVTLY 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 NQKFKKATLVTKSSATTAYMELRLSTSDSAVYYCARSTMTITNYMDYWGQCTSVTVSS 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 NQKFKKATLVTKSSATTAYMELRLSTSDSAVYYCARSTMTITNYMDYWGQCTSVTVSS 118
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

RESULT 15
GCB_MOUSE STANDARD; PRT; 336 AA.
ID GCB_MOUSE STANDARD; PRT; 336 AA.
AC P01866;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-2B chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (A ALLELE).
```

```
RX MEDLINE=80120716; PubMed=6766534;
RA Yamawaki-Kataoka Y., Kataoka T., Takahashi N., Obata M., Honjo T.;
RT "Complete nucleotide sequence of immunoglobulin gamma2b chain gene
RT cloned from newborn mouse DNA.";
RL Nature 283:786-789(1980).
RN [2]
RP SEQUENCE FROM N.A. (MPC 11).
RX MEDLINE=80081501; PubMed=117548;
RA Tucker P.W., Marcu K.B., Slightom J.L., Blattner F.R.;
RT "Structure of the constant and 3' untranslated regions of the murine
RT gamma 2b heavy chain messenger RNA.";
RL Science 206:1299-1303(1979).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=80081502; PubMed=117549;
RA Tucker P.W., Marcu K.B., Newell N., Richards J., Blattner F.R.;
RT "Sequence of the cloned gene for the constant region of murine gamma
RT 2b immunoglobulin heavy chain.";
RL Science 206:1303-1306(1979).
RN [4]
RP SEQUENCE FROM N.A. (B ALLELE).
RX MEDLINE=82173203; PubMed=6803173;
RA Ollo R., Rougeon F.;
RT "Mouse immunoglobulin allotypes: post-duplication divergence of gamma
RT 2a and gamma 2b chain genes.";
RL Nature 296:761-763(1982).
RN [5]
RP CARBOHYDRATE-LINKAGE SITE THR-105.
RX MEDLINE=94216359; PubMed=7512967;
RA Kim H., Yanaguchi Y., Masuda K., Matsunaga C., Yamamoto K.,
RA Irimura T., Takahashi N., Kato K., Arata Y.;
RT "O-glycosylation in hinge region of mouse immunoglobulin G2b.";
RL J. Biol. Chem. 269:12345-12350(1994).
CC -!- PTM: O-LINKED GLYCAN CONSISTS OF GAL-GALNAC DISACCHARIDE WHICH IS
CC MODIFIED WITH 2 SIALIC ACID RESIDUES.
CC -!- MISCELLANEOUS: DISULFIDE BONDS BY HOMOLOGY WITH OTHER IG GAMMA
CC CHAINS.
CC -!- MISCELLANEOUS: THE A ALLELE SEQUENCE IS SHOWN.
DR PIR: A02157; G2MS11.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00407; IGV; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 118 118 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 150 210
FT DISULFID 256 314
FT CARBOHYD 105 105
FT MOD_RES 336 336 REMOVED POST-TRANSLATIONALLY (PROBABLE).
FT VARIANT 163 163 Q -> R (IN B ALLELE).
FT VARIANT 194 194 T -> A (IN B ALLELE).
FT VARIANT 300 300 N -> D (IN B ALLELE).
FT VARIANT 301 301 M -> I (IN B ALLELE).
FT CONFLICT 25 25 L -> S (IN REF. 2 AND 3).
FT CONFLICT 36 36 S -> P (IN REF. 2 AND 3).
FT CONFLICT 239 239 I -> T (IN REF. 2 AND 3).
SQ SEQUENCE 336 AA; 36658 MW; 7D879662607C356E CRC64;

Query Match 13.6%; Score 479; DB 1; Length 336;
Best Local Similarity 27.6%; Pred. No. 9.2e-23;
Matches 151; Conservative 47; Mismatches 110; Indels 240; Gaps 14;

QY 121 AKTTPSPVYPLAPGSAQTNSMVTLCGLVKGYFPETVTVNWSGSLSSGVHTFPAVLQSD 180
```

```
Db 1 AKTTPPSVPLAPGCGDTGSSVTLGCLVKGYFPEVTVTNWNSGLSSVHTFPALQSG 60
Qy 181 LYTLSSSVTPSSWPSSTVTCNVAHPASSTKVKDKKIIVPRDGGPSEKSEINE-----KD 236
Db 61 LYTNSSSVTPSSWPSSTVTCNVAHPASSTVDKKLEP-----SGP---ISTINPCPKCE 114
Qy 237 LRK--KSELOGTALGNLKOIYYNSKAITSSSEKSDAQFLTNLTKGFTGHPWYNDLLV 294
Db 115 CHKCPAPNLEG---GPSVFIFPPNLIKDV-----LMI 142
Qy 295 DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINL 354
Db 143 SL-----TP----- 146
Qy 355 WIDGKQTTPIDKVKTSKKEVTVOELDLQARHYLHGKFLYNSDSFGKGKVGORGLIVFHSS 414
Db 147 ---KVTGVVD-----VSEDDPDVQISWVFN-----VEVHTA 176
Qy 415 EGSTVSYDLFDAQGOYPTDLLRIYRDNNTISSTLSISLYLTTYSIVMTQPTSLLSVAG 474
Db 177 QTQTHREDY-----NSTIRVVS----- 193
Qy 475 DRVTITCRASQSVNDVAVYQKQKPSKLLISYTSRRYAGVDPDRFSGSGYGTDFLTIS 534
Db 194 -----TLPIQ 198
Qy 535 SVQAEDAAVYFCQDYNPPTFGGKLEIKRADAAPTVSIFPPSSQELTSGGASVWCFL 594
Db 199 HDWNMSGREFKCKYNNKDLPSPIERTISKIGLYRAPQVYILPPPAEQLSRKDVSLTCLV 258
Qy 595 NNFYPKDINVKWKIDGSRQN-----GVLNSWTDDSDSYSSMSTLTITKDEYERHNS 649
Db 259 VGFNPGLDISVEWTSNGHTEENYKDTAPVLDL-----DGSYFYISKLNKMTSKWEKTD 311
Qy 650 YTCEATHK 657
Db 312 FSCNVRHE 319
RESULT 16
GCBM_MOUSE
ID GCBM_MOUSE STANDARD: PRT: 405 AA.
AC P01867;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE OF 335-405 FROM N.A.
RX MEDLINE=8222190; PubMed=6283537;
RA Yamawaki-Kataoka Y., Nakai S., Miyata T., Honjo T.;
RT "Nucleotide sequences of gene segments encoding membrane domains of
RT immunoglobulin gamma chains.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2623-2627(1982).
RN [2]
RP SEQUENCE OF 335-378 FROM N.A.
RX MEDLINE=8211295; PubMed=6799207;
RA Rogers J., Choi E., Souza L., Carter C., Word C.J., Kuehl M.,
RA Eisenberg D., Wall R.;
RT "Gene segments encoding transmembrane carboxyl termini of
RT immunoglobulin gamma chains.";
RL Cell 26:19-27(1981).
CC -!- ALTERNATIVE PRODUCTS: CELL LINES PRODUCING IGG CONTAIN TWO MRNA
CC SPECIES FOR IG GAMMA CHAINS. THE MAJOR SPECIES ENCODES SECRETED
CC GAMMA CHAINS. A LESS ABUNDANT SPECIES APPEARS TO ENCODE MEMBRANE-
CC BOUND CHAINS IN THAT IT CONTAINS AN ALTERNATIVE 3' END, ENCODED
CC IN SEPARATE EXONS, THAT IS HOMOLOGOUS WITH THE MEMBRANE-BOUND
CC SEGMENT OF MU CHAINS.
```

```
CC -!- MISCELLANEOUS: DISULFIDE BONDS BY HOMOMOLOGY WITH OTHER IG GAMMA
CC CHAINS.
CC -!- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 1-335 IS ASSUMED TO BE
CC IDENTICAL WITH THE CORRESPONDING REGION OF THE SECRETED FORM.
CC -!- MISCELLANEOUS: THE A ALLELE SEQUENCE IS SHOWN.
CC -----
CC This SWISS-PROT entry is by copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00462; AAB59659.1; ALT_INIT.
CC PIR: C02154; G2MSBM.
CC HSSP: P01842; 7FAB.
CC MGI: 96445; Igh-3.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003597; Ig_C1.
CC Pfam: PF00047; Ig_3.
CC SMART: SM00410; Ig_Like; 1.
CC SMART: SM00407; IG_C1; 2.
CC PROSITE: PS00290; IG_MHC; 1.
CC Immunoglobulin domain; Immunoglobulin C region; Transmembrane;
CC Alternative splicing.
CC NON_TER 1 1 INTERCHAIN (WITH A LIGHT CHAIN).
CC DISULFID 15 15
CC DISULFID 27 82
CC DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 118 118 INTERCHAIN (WITH A HEAVY CHAIN).
CC DISULFID 150 210
CC DISULFID 256 314
CC TRANSMEM 352 369
CC DOMAIN 370 405 POTENTIAL.
CC CYTOPLASMIC (POTENTIAL).
CC SEQUENCE 405 AA; 44330 MW; 89B3CFOA9B6D49FA CRC64;
Query Match 13.68; Score 479; DB 1; Length 405;
Best Local Similarity 27.68; Pred. No. 1.2e-22;
Matches 151; Conservative 47; Mismatches 110; Indels 240; Gaps 14;
Qy 121 AKTTPPSVPLAPGSAQAQTNMVTGLCLVKGYFPEVTVTNWNSGLSSVHTFPALQSD 180
Db 1 AKTTPPSVPLAPGCGDTGSSVTLGCLVKGYFPEVTVTNWNSGLSSVHTFPALQSG 60
Qy 181 LYTLSSSVTPSSWPSSTVTCNVAHPASSTKVKDKKIIVPRDGGPSEKSEINE-----KD 236
Db 61 LYTNSSSVTPSSWPSSTVTCNVAHPASSTVDKKLEP-----SGP---ISTINPCPKCE 114
Qy 237 LRK--KSELOGTALGNLKOIYYNSKAITSSSEKSDAQFLTNLTKGFTGHPWYNDLLV 294
Db 115 CHKCPAPNLEG---GPSVFIFPPNLIKDV-----LMI 142
Qy 295 DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINL 354
Db 143 SL-----TP----- 146
Qy 355 WIDGKQTTPIDKVKTSKKEVTVOELDLQARHYLHGKFLYNSDSFGKGKVGORGLIVFHSS 414
Db 147 ---KVTGVVD-----VSEDDPDVQISWVFN-----VEVHTA 176
Qy 415 EGSTVSYDLFDAQGOYPTDLLRIYRDNNTISSTLSISLYLTTYSIVMTQPTSLLSVAG 474
Db 177 QTQTHREDY-----NSTIRVVS----- 193
Qy 475 DRVTITCRASQSVNDVAVYQKQKPSKLLISYTSRRYAGVDPDRFSGSGYGTDFLTIS 534
Db 194 -----TLPIQ 198
Qy 535 SVQAEDAAVYFCQDYNPPTFGGKLEIKRADAAPTVSIFPPSSQELTSGGASVWCFL 594
```

```
Db 199 HQDWMGSEFKCKVNNKOLPSPERTISKIRGLVRAPQVILPPPAEQLSKRDVSLTCLV 258
Qy 595 NNFYPKDINVKWKIDGSRQN-----GVLNSWTDDSKDSTYSMSSTLTITKDEYERHNS 649
Db 259 VGFNPGDISVEMTSNGHTEENYKDTAPVLDS-----DGSFYIYKLNMTKSKWERTDS 311
Qy 650 YTCEATHK 657
Db 312 FSCNVRHE 319
RESULT 17
KACA_RAT
ID KACA_RAT STANDARD; PRT; 106 AA.
AC P01836;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain C region, A allele.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DA;
RA MEDLINE=82082587; PubMed=6273908;
RX Sheppard H.W., Gutman G.A.;
RT "Allelic forms of rat kappa chain genes: evidence for strong
RT selection at the level of nucleotide sequence.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:7064-7068(1981).
DR PIR: A02118; K1RTA.
DR HSSP: P01842; 2MCG.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00407; IGcl; 1.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DISULFID 26 86
FT DISULFID 106 106
SQ SEQUENCE 106 AA; 11732 MW; B7E120D970DDDD66 CRC64;
Query Match 13.4%; Score 472; DB 1; Length 106;
Best Local Similarity 83.8%; Pred. No. 5.4e-23;
Matches 88; Conservative 11; Mismatches 6; Indels 0; Gaps 0;
Qy 567 ADAAPTTSIFPPSPBQLTSGGASVYCFLNFPKIDINVKWKIDGSRQGVLSNWTDDSD 626
Db 1 ADAAPTTSIFPPSPBQLTSGGATVYCFVNNFPRDISVKWKIDGSEQRDGVLSNWTDDSD 60
Qy 627 KDSTYSMSSTLTITKDEYERHNSYTCETHKTSPIVKSFNRRNE 671
Db 61 KDSTYSMSSTLTITKVEYERHNLTYTCEVYVHKTSSSPVYKSFNRRNE 105
RESULT 18
GCAB_MOUSE
ID GCAB_MOUSE STANDARD; PRT; 335 AA.
AC P01864;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2A chain C region, B allele.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
```

```
RX MEDLINE=82037861; PubMed=6170065;
RA Schreier P.H., Bothwell A.L.M., Mueller-Hill B., Baltimore D.;
RT "Multiple differences between the nucleic acid sequences of the
RT IgG2aa and IgG2ab alleles of the mouse";
RL Proc. Natl. Acad. Sci. U.S.A. 78:4495-4499(1981).
RN [2]
RP SEQUENCE.
RX MEDLINE=8203777; PubMed=6794027;
RA Dognin M.J., Lauwereys M., Strosberg A.D.;
RT "Multiple amino acid substitutions between murine gamma 2a heavy
RT chain FC regions of Ig1a and Ig1b allotypic forms.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:4031-4035(1981).
CC -!- MISCELLANEOUS: THE SEQUENCE DIFFERS FROM THAT OF THE A ALLELE,
CC FROM BALB/C MICE, AT 15% OF THE POSITIONS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00479; -; NOT_ANNOTATED_CDS.
DR PIR: A02153; G2MSAB.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00407; IGcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
SQ SEQUENCE 335 AA; 36596 MW; FA3382792CBB13C6 CRC64;
Query Match 13.3%; Score 467; DB 1; Length 335;
Best Local Similarity 25.5%; Pred. No. 5.1e-22;
Matches 149; Conservative 40; Mismatches 110; Indels 286; Gaps 13;
Qy 121 AKTTPSVYPLAPGSAAGTNSMTVLGCLVKGYFPEPVTVTVNSGSLSSGVHFFAVLQSD 180
Db 1 AKTAPSVYPLVPVCGGTGSGTGLCLVKGYFPEPVTLTVNNSGSLSSGVHFFAVLQSG 60
Qy 181 LYTLSSTVTPSSTWPSSTVTCNVAHPASSTKVDKIVPRDSGGPSEKSEINEKDLRKK 240
Db 61 LYTLSSTVTPSSTWPSSTVTCNVAHPASSTKVDKIEPR----- 100
Qy 241 SELQGTALGNLKOIYVYNSKAITSSKSAADQFLNTLLFKGFTTGHWPYNDLLVDLGSTA 300
Db 101 ----- 100
Qy 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMGVGLTLDNNRLTEKKVYINLWIDGKQ 360
Db 101 ----- 106
Qy 361 TTVPIDKVKTSKKEVTYQELDLQARHYLHGKGLYNSDFGKVGQGLIVFHSSEGSTVS 420
Db 107 PCPPHQVRVPCAA----- 126
Qy 421 YDLFDAOGQYPTLLRIYRDNTTSSLSLSLYLYTTTSIVMTQTPTSLVLSAGDRVTIT 480
Db 127 -SVFIFFPKIKDVLN----- 152
Qy 481 CKASQSVSNDAVWYQOKPGQSPKLLIYSTSSRYAGVDPDRFSGSGYGTDTITISVQAED 540
Db 153 ----- 174
Qy 541 AAVFYCOQDYNSPPTFGGTTKLEIKRAD----- 569
Db 175 AQTQTHREDYNS--TLRVVSALPIQHODMWSGKEFKCKVNNRNLPSPIEKTISKPRGVR 232
```


or send an email to license@isb-sib.ch).

```
CC EMBL: J00539; AAA38172.1; -.
CC PIR: A02038; G2MS43.
DR HSSP: P01810; 2FBJ.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig; 1.
DR SMART: SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 137 IG HEAVY CHAIN V REGION S43.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 20 49 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 50 54 FRAMEWORK-2.
FT DOMAIN 55 68 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 69 85 FRAMEWORK-3.
FT DOMAIN 86 117 FRAMEWORK-4.
FT DOMAIN 118 122 D SEGMENT.
FT DOMAIN 123 137 JH2 SEGMENT.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 137 137
SQ SEQUENCE 137 AA; 15200 MW; ADD5881BF44B8EC9 CRC64;

Query Match 12.9%; Score 454; DB 1; Length 137;
Best Local Similarity 70.0%; Pred. No. 9.8e-22;
Matches 84; Conservative 15; Mismatches 19; Indels 2; Gaps 1;

Qy 1 EVQLQQSGPDLVKGASVKISKASGYSTGYTHHWVKQSPGKGLGWIRNPNGVTLY 60
Db 20 QVQLQQSGAEFLVRGSSVKMSCKASGYTFTSYGYNWVKRPGQGLEWIGYINPGNGYIN 79
Qy 61 NQKFKDKATLVDRKSSSTAYMELRSLSSEDSAVYYCARSTMI-TNYVMDYWGQGTSTVVS 120
Db 80 NEHFRSKATLVDRKSSSTAYMELRSLSSEDSAVYYCARVRL--GRYFDYWGQGTSTVVS 137

RESULT 22
HV02_MOUSE
ID HV02_MOUSE STANDARD; PRT; 140 AA.
AC P01746;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 93G7 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/J;
RX MEDLINE=82152818; PubMed=6801765;
RA Sims J., Rabbits T.H., Estess P., Slaughter C., Tucker P.W.,
RA Capra J.D.;
RT "Somatic mutation in genes for the variable portion of the
RT immunoglobulin heavy chain.";
RL Science 216:309-311(1982).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
```

```
KW Immunoglobulin V region; Antiarsonate antibody; Hybridoma; Signal.
FT SIGNAL 1 19
FT CHAIN 20 140 IG HEAVY CHAIN V REGION 93G7.
FT NON_TER 140 140
SQ SEQUENCE 140 AA; 15514 MW; 25A4CBBE31DA5CE8 CRC64;

Query Match 12.8%; Score 450.5; DB 1; Length 140;
Best Local Similarity 71.9%; Pred. No. 1.7e-21;
Matches 87; Conservative 14; Mismatches 19; Indels 1; Gaps 1;

Qy 1 EVQLQQSGPDLVKGASVKISKASGYSTGYTHHWVKQSPGKGLGWIRNPNGVTLY 60
Db 20 EVQLQQSGAEFLVRGSSVKMSCKASGYTFTSYGYNWVKRPGQGLEWIGYINPGNGYIN 79
Qy 61 NQKFKDKATLVDRKSSSTAYMELRSLSSEDSAVYYCARSTMI-TNYVMDYWGQGTSTVVS 119
Db 80 NEKFKGKTLTVDRKSSSTAYMELRSLSSEDSAVYYCARSHYGGSYDFDYWGQGTPLTVS 139
Qy 120 S 120
Db 140 S 140

RESULT 23
GCB_RAT
ID GCB_RAT STANDARD; PRT; 333 AA.
AC P20761;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Brueggemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR: PS0018; PS0018.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_like.
DR Pfam: PF00047; Ig; 3.
DR SMART: SM00410; IG_like; 1.
DR SMART: SM00407; IGcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA; 36497 MW; 55F8B64D48D460A6 CRC64;

Query Match 12.7%; Score 446; DB 1; Length 333;
Best Local Similarity 25.3%; Pred. No. 1e-20;
Matches 140; Conservative 54; Mismatches 131; Indels 228; Gaps 14;

Qy 121 AKTTPPSVYPLAPGSAQTNSMVTLCGLVKGYFPPEVTVTWNSGLSSGVHTFFAVLQSD 180
Db 1 AQTTPASVYPLAPGCGDFTSTVTLCGLVKGYFPPEVTVTWNSGALSSDVHTFFAVLQSG 60
Qy 181 LYTLLSSSVTVPSSTWPSSTVTCNVNAHPASSTKVDKKIVPRDSSGGPSEKSEINEKLRKK 240
Db 61 LYTLLTSSVT--SSTWPSQTVTCNVNAHPASSTKVDKKVKKERRNG----- 101
```

```
QY 241 SELQ TALGNLKIYYNKAITSSEKSAQDLNTLLEKGFTHGPHWYNLDLLDGLSTA 300
Db 102 -----
QY 301 ATSEYEGSSVDLYGAYGYOCAGGTPNKACMYGGVTLHDNNRLTEEEKVPLNLWDGKO 360
Db 102 -----
QY 361 TTVPIDKVKTSKEVTYQELDLOARHLHGKFGCLYNSDSFGKQVORGLIVFHSSEGSTVS 420
Db 126 VFT -----
QY 421 YDLFDAQQGYPTDLLRIYRDNNTTISLSLSISLYLTTSIVMTQTPTSLLSAGDRVT-- 478
Db 147 CVVDVSEEPDQVQSFVNNVEVHTAQTPREEQYNSTF---RVVSALPIQHQQHMSCK 203
QY 479 -ITCKASQSVSNDAVWYQKQPGSKLLSYSSRYAGVDPDRFSGSGYGTDTFTLTSSVO 537
Db 204 EFKCKVN-----NKALPSP---IEKTSKPKGLVRK----- 231
QY 538 AEDAAVYFCQODYNSPPTFGGKLEIKRADAAPTYSIFPPSSEQLTSGGASVVCFLNMF 597
Db 232 -----
QY 598 YPRDINVKWKIDGSRQNGVLNSWTDQDSKSTYSMSSTLTTLTKDEYERHNSYTCETHK 657
Db 259 LPNDIGVETNSGHIEKN-YKNTPEWMSD-DGSFFMYSKLNVERSWDSRAPFVCSVWHE 316
QY 658 -TSTSPVKSFRN 659
Db 317 GLHNHHRVKSISR 329

RESULT 24
HV48_MOUSE STANDARD; PRT; 138 AA.
AC P03980;
DT 23-OCT-1986 (Rel. 02, Created)
DT 23-OCT-1986 (Rel. 02, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain v region TEPC 1017 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84248078; PubMed=6429663;
RA Gilliam A.C., Shen A., Richards J.E., Blattner F.R., Mushinski J.F.,
RA Tucker P.W.;
RT "Illegitimate recombination generates a class switch from C mu to C
RT delta in an IgD-secreting plasmacytoma."
RL Proc. Natl. Acad. Sci. U.S.A. 81:4164-4168(1984).
DR PIR; A02033; HVMST7.
DR HSSP; P01810; 2FB3.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
KW Immunoglobulin v region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 138 IG HEAVY CHAIN V REGION TEPC 1017.
FT DOMAIN 21 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DOMAIN 118 127 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 128 138 FRAMEWORK-4.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 138 138
SQ SEQUENCE 138 AA; 15576 MW; 748157E4C6907B8E CRC64;
```

```
Query Match 12.4%; Score 437.5; DB 1; Length 138;
Best Local Similarity 68.3%; Pred. No. 1e-20;
Matches 82; Conservative 20; Mismatches 17; Indels 1; Gaps 1;

QY 1 EVLOQSGDPLVKPGASVKISKASGYSTGYMHWKSPGKGLWIGRIINPNNGVTLY 60
Db 20 QVLOQPGAELVKPGASVQLSKASGHTFTNYIHWKQRPQGLWIEINPNNGRSNY 79

QY 61 NQKFKDKATLTVDKSTTAYMELRSLTSEDSAVYYCARSTMTNYYVMDYWGQGSTVTVSS 120
Db 80 NEKFNKATLTVDKSSSTAYMQLSLTPEEFAVYYCARSDGYDMFV-YWGQSTLTFTSA 138

RESULT 25
GCC_RAT STANDARD; PRT; 329 AA.
ID GCC_RAT
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2c chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166903; PubMed=3127222;
RA Brueggemann M., Delmastro-Galfrè P., Waldmann H., Calabi F.;
RT "Sequence of a rat immunoglobulin gamma 2c heavy chain constant
RT region cDNA: extensive homology to mouse gamma 3."
RL Eur. J. Immunol. 18:317-319(1988).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X07189; CAA30169.1; -.
DR PIR; S00847; S00847.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00410; IG-like; 1.
DR SMART; SM00407; IGcl; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 222 CH2.
FT DOMAIN 223 329 CH3.
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 82 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 111 111 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 113 113 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 143 203
FT DISULFID 249 307
SQ SEQUENCE 329 AA; 36571 MW; 5FCD7B7933850773 CRC64;

Query Match 12.4%; Score 437.5; DB 1; Length 329;
Best Local Similarity 26.1%; Pred. No. 3.3e-20;
Matches 140; Conservative 47; Mismatches 125; Indels 225; Gaps 16;

QY 121 AKTTPPSVYPLAPGSAAQTNNSVMTGLCLVKGYFPEPVTVTVNSGSLSSGVHTFPAVLQSD 180
Db 1 ARTTAPSVYPLVPGSGTSGSLVTLGCLVKGYFPEPVTVTVNSGSLSSGVHTFPAVLQSG 60
```

```

Db      1 QVQLQPGTELVKGASVNLISCKASGYFTFSYWHHWJTRQRPCQGLEWIGGINPSNGGTNY 60
Qy      61 NQKFKDKATLVDRKSSSTAYMELRSLTSEDSAVYYCARSMITNVMVDYMGQGTSVTSS 120
        I::I ::::::::::: II:: : ::::::::::::::: : II::II::II::II::
Db      61 NEKFKSKATLVDRKSSSATYMQLSPTPTSSEDSAVYYCARWDYEGDRIYEDVWGCTGTTVTSS 120

RESULT 27
HV09_MOUSE
ID   HV09_MOUSE          STANDARD;              PRT;       117 AA.
AC   P01753; P11271;
DT   21-JUL-1986 (Rel. 01, Created)
DI   01-JUL-1989 (Rel. 11, Last sequence update)
DE   15-JUL-1999 (Rel. 38, Last annotation update)
OS   Ig heavy chain v region 186-1 precursor.
OC   Mus musculus (Mouse).
NC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX   NCBI_TaxID=10090;
[1]
SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6;
RA   MEDLINE=81234548; PubMed=6789376;
RA   Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA   Baltimore D.;
RT   "Heavy chain variable region contribution to the NPb family of
RL   antibodies: somatic mutation evident in a gamma 2a variable region."
RL   Cell 24:625-637(1981).
CC   -I- MISCELLANEOUS: THIS GERM LINE GENE BELONGS TO A SET OF CLOSELY
        RELATED GENES THAT COULD ENCODE V REGIONS OF NPB ANTIBODIES.
PTR; B02034; HYMS61.
DR   HSP; P01810; 2FBJ.
DR   InterPro: IPR003006; Ig_MHC.
DR   InterPro: IPR003596; Ig_v.
DR   Pfam; PF00047; Ig; 1..
DR   SMART; SM00406; IGV; 1..
DR   Immunoglobulin V region; Signal.
FT   SIGNAL             1
FT   CHAIN              20 117
FT   DOMAIN             20 49
FT   DOMAIN             50 54
FT   DOMAIN             55 68
FT   DOMAIN             85 85
FT   DOMAIN             86 117
FT   DISULFID           41 115
FT   NON_TER            117 117
SQ   SEQUENCE 117 AA; 12890 MW; 16191A088CB17F5A CRC64;

Query Match               12.1%; Score 425; DB 1; Length 117;
Best Local Similarity    79.6%; Pred. No. 4.9e+20;
Matches                  78; Conservative 11; Mismatches 9; Indels 0; Gaps

Qy      1 EVQLQQSGPDLVKGASVKISKASGYSFTGYIMHWWKQSPGKLEWIGRINPNNGVTLY 60
        :IIII I :::::::::::IIIIII::II::IIIIII::II::IIIIII::II::II
Db      20 QVQLQPGAEVLVKGASVKLSCASGYFTFSYWHHWKQRPGRGLEWIGRIDPNSGGTKY 79

Qy      61 NQKFKDKATLVDRKSSSTAYMELRSLTSEDSAVYYCAR 98
        I::I :::::::::::IIIIII::IIIIII::IIIIII::IIIIII::IIIIII

Db      80 NEKFKSKATLVDRKSSSATYMQLSPTPTSSEDSAVYYCAR 117

RESULT 28
KV5A_MOUSE
ID   KV5A_MOUSE          STANDARD;              PRT;       149 AA.
AC   P01633;
DT   21-JUL-1986 (Rel. 01, Created)
DI   21-JUL-1986 (Rel. 01, Last sequence update)
DE   15-JUL-1999 (Rel. 38, Last annotation update)
OS   Ig kappa chain v-v region MPC11 precursor.
OC   Mus musculus (Mouse).
NC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RX   NCBI_TaxID=10090;
```



```
Db 80 NEKFSKATLTVDKPSSTAYMQLSSTSEDSAVYYCAR 117
RESULT 33
HV49_MOUSE STANDARD; PRT; 117 AA.
AC P06328;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region VH558 B4 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099340; PubMed=2578321;
RA Yancopoulos G.D., Alt F.W.;
RT "Developmentally controlled and tissue-specific expression of
RT unrearranged VH gene segments.";
RL Cell 40:271-281(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M13788; AAA38506.1; -.
DR PIR; A02035; MHMSB4.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION VH558 B4.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12834 MW; B8862FAC67ABD345 CRC64;

Query Match 11.8%; Score 414; DB 1; Length 117;
Best Local Similarity 78.4%; Pred. No. 2.4e-19;
Matches 76; Conservative 10; Mismatches 11; Indels 0; Gaps 0;

Qy 2 VOLQSGDPLVKPGASVKISCKASGYSTFGYMHVWKSPGKGLWGIRPNNGVTLYN 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 21 VOLQPGAEELVKPGASVKISCKASGYSTFGYMHVWKSPGKGLWGIRPNNGVTLYN 80
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Qy 62 QKFKDKATLTVDKSSSTAYMELSLTSEDSAVYYCAR 98
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 81 ERKFSKATLTVDKPSSTAYMQLSSTSEDSAVYYCTR 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |

RESULT 34
HV06_MOUSE STANDARD; PRT; 117 AA.
ID HV05_MOUSE
AC P01750;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region 102 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=81234548; PubMed=6788376;
RA Bothwell A.L.M., Paskind M., Reth M., Imanishi-Kari T., Rajewsky K.,
RA Baltimore D.;
RT "Heavy chain variable region contribution to the NPb family of
RT antibodies: somatic mutation evident in a gamma 2a variable region.";
RL Cell 24:625-637(1981).
CC -!- MISCELLANEOUS: THIS GERMLINE GENE BELONGS TO A SET OF CLOSELY
CC RELATED GENES THAT COULD ENCODE V REGIONS OF NPb ANTIBODIES.
DR PIR; A02032; HVMS02.
DR HSSP; P01810; 2FBJ.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; Igv; 1.
DR Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117 IG HEAVY CHAIN V REGION 102.
FT DOMAIN 20 49 FRAMEWORK-1.
FT DOMAIN 50 54 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 55 68 FRAMEWORK-2.
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12867 MW; 740A65DD851FCA8C CRC64;

Query Match 11.7%; Score 413; DB 1; Length 117;
Best Local Similarity 80.2%; Pred. No. 2.7e-19;
Matches 77; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

Qy 2 VOLQSGDPLVKPGASVKISCKASGYSTFGYMHVWKSPGKGLWGIRPNNGVTLYN 61
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 21 VOLQPGAEELVKPGASVKISCKASGYSTFGYMHVWKSPGKGLWGIRPNNGVTLYN 80
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Qy 62 QKFKDKATLTVDKSSSTAYMELSLTSEDSAVYYCA 97
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |
Db 81 QKFKDKATLTVDKSSSTAYMQLSSTSEDSAVYYCA 116
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | |

RESULT 35
GC2_CAVPO STANDARD; PRT; 329 AA.
ID GC2_CAVPO
AC P01862;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-2 chain C region.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE OF 1-3.
RA Trischmann T.M.;
RL Submitted (APR-1975) to the PIR data bank.
RL [2]
RP SEQUENCE OF 4-68.
RX MEDLINE=71058471; PubMed=5538606;
RA Birshtein B.K., Hussain Q.Z., Cebrera J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
RT immunoglobulin-G(2). 3. Amino acid sequence of the region around the
RT half-cystine joining heavy and light chains.";
RL Biochemistry 10:18-25(1971).
RN [3]
RP SEQUENCE OF 69-133 AND 312-329.
RX MEDLINE=71058486; PubMed=5538616;
RA Turner K.J., Cebrera J.J.;
RT "Structure of heavy chain from strain 13 guinea pig
```

```
RT immunoglobulin-G(2). II. Amino acid sequence of the carboxyl-terminal
RL and hinge region cyanogen bromide fragments."
RL Biochemistry 10:9-17(1971).
RN [4]
RP SEQUENCE OF 134-226.
RX MEDLINE=75036072; PubMed=4429665;
RA Tracey D.E., Cebra J.J.;
RT "Primary structure of the CH2 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4796-4803(1974).
RN [5]
RP SEQUENCE OF 227-311.
RX MEDLINE=75036073; PubMed=4609467;
RA Trischmann T.M., Cebra J.J.;
RT "Primary structure of the CH3 homology region from guinea pig IgG2
RT antibodies.";
RL Biochemistry 13:4804-4811(1974).
RN [6]
RP DISULFIDE BONDS.
RX MEDLINE=71058474; PubMed=4922544;
RA Oliveira B., Lamm M.E.;
RT "Interchain disulfide bridges of guinea pig gamma-2-immunoglobulin.";
RL Biochemistry 10:46-51(1971).
CC -!- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM POOLED SERUM OF STRAIN
CC 13 INBRED GUINEA PIGS.
DR PIR: A02151; G2GP.
DR HSP: P01842; 7FAB.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_cl.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig; 2.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00407; IGCL; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein.
FT NON_TER 1
FT DISULFID 16 16 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 28 79
FT DISULFID 105 105 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 110 110 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 142 202 N-LINKED (GLCNAC. . .).
FT CARBOHYD 178 178
FT DISULFID 248 308
SQ SEQUENCE 329 AA; 36074 MW; 5D231B7164D1FBA9 CRC64;

Query Match 11.7%; Score 413; DB 1; Length 329;
Best Local Similarity 23.6%; Pred. No. 1.le-18;
Matches 136; Conservative 61; Mismatches 104; Indels 276; Gaps 17;

QY 120 SAKTTPPSVYPLAGSAAQTNSMVTGLGVGYFPEPTVTWNSSGSLSSGVHTFPAVLQS 179
DB 1 SARTTAPSVFPLAASCVDTSGSMVTGLGVGYFPEPTVTWNSSGSLSSGVHTFPAVLQS 60

QY 180 DLYLTSSVTPSSWTPSETVTCNVAHPASSTKVDKKI-----VPRDS 222
DB 61 GLYLSMTMVTPS-----SQKATCNVAHPASSTKVDKVTPEIRTPZPBCTCPKCPPENL 116

QY 223 GPPSEKSEINEKDKRKSELOGTALGNLKQIYYNSKAITSSKSAOQFLTNLLFKGF 282
DB 117 GGPS-----VFIPPKP----- 128

QY 283 FTGHPWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTGPNKTCMAGGVVTLHDNN 342
DB 129 -----KDTLMISL-----TPRVT-CVVVDVS----- 148

QY 343 RUTEKKVPINLWIDGK-----QTTVPIDKVKTSKKEVTQVQELDLQARHYLHGKFGLYNS 397
DB 149 --QDEPEVQFTWFVDNKPVGNAETPRVEQYNTTFRVES--LPQHODWLGRK----- 198

QY 398 DFGGKVGORGLIVFHSSEGSTVSYDLFDAQGOYPTLLRIYRDNNTTSSTSLSISLYLT 457
DB 199 ----- 198
```

```
QY 458 TSIVMTQTPTSLLSAGDRVTITCKASQSVNDVAMYQOKPGQSPKLLISYTSRRYAGVP 517
DB 199 -----EFKCKV-----YNKALPAP---IEKTISKGAP 224

QY 518 DRFSGSGYGTDFTLTSSVQAEADAAVYFCQDYNPPFTFGGTKLEIKRADAAPTVSIFP 577
DB 225 -----RMPDVYTLTP 233

QY 578 PSSEOLTSGGASVVCFLNNFYPKDINVKW---KIDGSRQNGVLNSWTDQDSKDTYSMS 634
DB 234 PSRDELSKSKSVTCLIIINFFPADIHVEWASNRVPSVSEK--YKNTPTPIEDA-DGSYELY 290

QY 635 STLTLTKDEYERHNSYTCETATHTKSTSPIV-KSPNRN 670
DB 291 SKLTVDKSAWDOGTVTYTCVHVEALHNHVTOKAISRS 327

RESULT 36
KV4B_HUMAN
ID KV4B_HUMAN STANDARD; PRT; 133 AA.
AC P06313;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region JI precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86041853; PubMed=2997712;
RA Klobeck H.G., Bornkamm G.W., Combriato G., Mocikat R., Pohlenz H.D.,
RA Zachau H.G.;
RT "Subgroup IV of human immunoglobulin K light chains is encoded by a
RT single germline gene."
RL Nucleic Acids Res. 13:6515-6529(1985).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z00022; CAA77317.1; -.
DR PIR; A01904; K4HUJI.
DR HSSP; P80362; 1WTL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IgV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 20
FT CHAIN 21 133 IG KAPPA CHAIN V-IV REGION JI.
FT DOMAIN 21 43 FRAMEWORK-1.
FT DOMAIN 44 60 COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 61 75 FRAMEWORK-2.
FT DOMAIN 76 82 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 83 114 FRAMEWORK-3.
FT DOMAIN 115 122 COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 123 132 FRAMEWORK-4.
FT DISULFID 43 114 BY SIMILARITY.
FT NON_TER 133 133
SQ SEQUENCE 133 AA; 14632 MW; 5FB3953066744AF4 CRC64;

Query Match 11.5%; Score 406.5; DB 1; Length 133;
Best Local Similarity 63.9%; Pred. No. 8.2e-19;
Matches 85; Conservative 16; Mismatches 21; Indels 11; Gaps 3;

QY 444 ISSTSLSISLYLTT-----SIVMTQTPTSLLSVAGDRVTITCKASQSV-----SNDVAV 493
```



```

Db      : | : ||| |: :   |||||: || || :-| || |||||: | :||
2 VLOTQVPSILLWISGAYGOIVMTQSPDLSVLASGLGRATINCKSSOSILYSSDNKNYLAW 61

Qy      494 YQKPGQSPKLLSYSTSYRAGVPDRFSGSGYGTFTLTISYYQAEDAAVFYFCQQDYNSP 553
        ||||||| ||||| : | : ||||||| ||||| ||||| ||||| ||||| ||||| :| :
62 YQKPGQPPKLLIYWASTRESGVDPDRFSGSGGTDTFTLTISSLQAEADVAVYCCQ-YDTI 120

Qy      554 PTFGGGTGLEIKR 566
        |||||||: |||||
Db      121 PTFGGGTKEIKR 133

RESULT 37
KV4C_HUMAN STANDARD;          PRS;    134 AA.
ID     KV4C_HUMAN AC
AC     P06314;
DT     01-JAN-1988 (Rel. 06, Created)
DT     01-APR-1988 (Rel. 07, Last sequence update)
DT     15-JUL-1999 (Rel. 38, Last annotation update)
DE     Ig kappa chain V-IV region B17 precursor.
OS     Homo sapiens (human).
OC     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN     [1]
RP     SEQUENCE FROM N.A.
RX     MEDLINE=86041854; PubMed=2997713;
RA     Marsh P., Mills F., Gould H.;
RT     "Detection of a unique human V kappa IV germline gene by a cloned
RT     cDNA probe.";
RL     Nucleic Acids Res. 13:6531-6544(1985).
RN     [2]
RP     REVISION TO 76.
RA     Marsh P.;
RL     Submitted (OCT-1986) to the EMBL/GenBank/DDBJ databases.
CC     This SWISS-PROT entry is copyright. It is produced through a collabora-
CC     between the Swiss Institute of Bioinformatics and the EMBL outpost
CC     the European Bioinformatics Institute. There are no restrictions on
CC     use by non-profit institutions as long as its content is in no way
CC     modified and this statement is not removed. Usage by and for comm-
CC     entities requires a license agreement (See http://www.isb-sib.ch/anno-
CC     or send an email to license@isb-sib.ch).
-----
DR      ENBL: X02990; CAA26733.1; -.
DR      PIR: A01905; K4HU17.
DR      HSSP: P80362; IWLTL.
DR      InterPro: IPR003006; Ig_MHC.
DR      InterPro: IPR003596; Ig_V.
DR      Pfam: PF000047; Ig_1.
DR      SMART: SM00406; IGV; 1.
KW      immunoglobulin v region; Signal.
FT      SIGNAL             1..20
FT      CHAIN              21..134
FT      DOMAIN              21..43
FT      DOMAIN              44..60
FT      DOMAIN              61..75
FT      DOMAIN              76..82
FT      DOMAIN              83..114
FT      DOMAIN              115..121
FT      DOMAIN              122..133
FT      DISULFID            43..114
FT      NON_TER             134..134
SQ      SEQUENCE          134 AA; 14966 MW; 6413A22FD0738832 CRC64;
Query Match           11.5%; Score 406; DB 1; Length 134;
Best Local Similarity 63.2%; Pred. No. 8.9e-19;
Matches 84; Conservative 15; Mismatches 24; Indels 10; Gaps

Qy      444 ISSTSLSISLYVT-----SIWMTQPTSLVSAGRVTITCKASOV-----SNDVAM 493
        : | : ||| :| :   |||||: || || :-| || |||||: | :||
Db      2 VLOTQVPSILLWISGAYGOIVMTQSPDLSVLASGLGRATINCKSSOSILYSSDNKNYLAW 61

```

```
RESULT 39
GC3_MOUSE STANDARD; PRT; 329 AA.
AC P22436;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region, secreted form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Weis J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; ; NOT_ANNOTATED_CDS.
DR PIR; B02156; G3MSC.
DR HSP; P01857; 1FCL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-CL.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IGCL; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
SQ SEQUENCE 329 AA; 36228 MW; F45827174182BAD6 CRC64;

Query Match 11.4%; Score 401; DB 1; Length 329;
Best Local Similarity 24.6%; Pred. No. 6e-18;
Matches 135; Conservative 51; Mismatches 111; Indels 252; Gaps 15;

QY 123 TTPPSVYPLAPGSAATNSMVLGCLVKGYFPEPVVTVVNSGSLSGVITFPAVLQSDLY 182
Db ||||| ||| : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2 TTPASVYPLVPGCDSTSGSVTLGCLVKGYFPEPVTVKWNYGALSSGVRTSVSLQSGFY 61
QY 183 TLSSSVTPSSTPSTVTCNVAHPASSTKVDKKIVPRDSGGGSEKSEINEKDLRKKSE 242
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 62 SLSSLVTPSSVTPSTVTCNVAHPASKTELKRIEPR-----IPKPS 105
QY 243 LGTAL--GNLKOIYYNSKAITTSSEKSAQDLTNTLLEKFGFTGHPWYNLLVDLGSTA 300
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 106 PGSSCCPPGNI-----LG--- 118
QY 301 ATSEYEGSSVDLYGAYGYQCAGGTGNKTYGCVGVTLHDNNRLTEEKVPTNLWTDGQK 360
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 119 -----GPSVFIP-----PPK-----PKDALMSLTLPKV 141
QY 361 TTVPIDKVTSKKEVTVQELDLQARHYLHGKFLYNSDSFGKVGQRLIVFHSSEGSTVS 420
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
QY 142 TCVVVD-----VSEDD----- 152
QY 421 YDLFDAQGOYPTDLLRIYRDNTTISSTLSLSILYLTTSIVMTQTPTSLVSAGDRVTT 480
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||

Db 153 -----PDVHVSWEVDNKEVHT----- 168
QY 481 CKASQSVNDVAWYQKQPCQSPKLLISYTSRYAGVDPDRFSGSGYGTDF-TLTISVQAE 539
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 169 -----AWTOPREAQ-----YNSTRFVVSALPIQHQ 193
QY 540 D---AAVYFCQODYNSPPTFGGKLEIKRADAAAPTVSIFPPSSEOLTSGGASVWCFLAN 596
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 194 DMWRGKEFKCKVNNKALPAPIERTISKPKRAQTPQVYTIPTPPREOMSKKVSLTCLVTN 253
QY 597 FYPKIDINVKWKIDGSRQGVNSWTFDQDSK-----DSTVSMSTLTLTQKDEYERHN 648
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 254 FFESEALSVEWE-----RNGEL-----EODYKNTPTILDSDGYFLYSLKLTVDTDSWLQGE 303
QY 649 SYTCEATHK 657
Db ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 304 IFTCSVVEH 312

RESULT 40
GC3M_MOUSE STANDARD; PRT; 398 AA.
AC P03987;
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6092053;
RA Weis J.A., Word C.J., Rimm D., Der-Balan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
RL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers J., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene
segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00451; AAB59655.1; -
DR EMBL; V01526; CA24767.1; ALT_SEQ.
DR PIR; A02155; G3MSM.
DR HSP; P01857; 1FCL.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-CL.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IGCL; 2.
DR PROSITE; PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 223 CH2.
FT DOMAIN 224 327 CH3.
```



```
Query Match 11.3%; Score 396.5; DB 1; Length 121;
Best Local Similarity 62.8%; Pred. No. 3e-18;
Matches 76; Conservative 16; Mismatches 28; Indels 1; Gaps 1;

QY 1 EVLOQSGDPLVKPGASVKISCKASYFTGYMHVWVKSPGKLEWIGRINPNNGVTLY 60
Db 1 EAQLOQSGAELVRPGTSVSKISKAAGYFTNYWGWKRPBGHLEWIGDIYPGGFTNY 60
QY 61 NOKFKDKATLTVDKSSATTAYMELRSLTSDSDSAVIYCARSTMI--TNVYMDYWGQGTSTVTVS 119
Db 61 NDLKAGKATLTADTSSSTAYIQLSLTSDSDSAIYHCARGIYNNSSPYFDSWGQGTTLTVS 120
QY 120 S 120
Db 121 S 121

RESULT 43
GC4_HUMAN STANDARD; PRT; 327 AA.
AC P01861;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-4 chain C region.
GN IGHG4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83157104; PubMed=6299662;
RA Ellison J.W., Buxbaum J.N., Hood L.E.;
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
RL DNA 1:11-18(1981).
RN [2]
RP SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE=70207560; PubMed=4192699;
RA Pink J.R.L., Buttery S.H., de Vries G.M., Milstein C.;
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
constant region of a gamma 4 chain.";
RL Biochem. J. 117:33-47(1970).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; K01316; AAB59394.1; ALT_INIT.
DR PIR; A02150; G4HU.
DR HSP; P01842; 7FAB.
DR Genew; HGNC:5528; IGHG4.
DR MIM; 147130; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig-C1.
DR InterPro; IPR003600; Ig_Like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; IG_Like; 1.
DR SMART; SM00407; IGc1; 2.
DR PROSITE; PS00290; IG_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 98 CH1..
FT DOMAIN 99 110 HINGE.
FT DOMAIN 111 220 CH2.
FT DOMAIN 221 327 CH3.
FT DISULFID 14 14 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
```

```
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 141 201
FT DISULFID 247 305
SQ SEQUENCE 327 AA; 35940 MW; 3EBDB811EF208E7A CRC64;

Query Match 11.3%; Score 396.5; DB 1; Length 327;
Best Local Similarity 23.3%; Pred. No. 1.1e-17;
Matches 131; Conservative 52; Mismatches 102; Indels 27; Gaps 16;

QY 121 AKTTPPSVYPLAPGSAQTNSMVTLCGLVKGYFPEPTVTWNSGSLSSGVHFPAYLQSD 180
Db 1 ASTKGPSVPLAPCSRSTSESTAALGCLVKDYFPEPTVTVSNWGALTSGVHTFPAYLQSS 60
QY 181 -LYTSSSVTPSPSTWPTVCNVAHPASSTKVKDKKI-----VPRDSGGPSE 227
Db 61 GLYSLSVVTPSSSLGTCTYCNVDHKPSNTKVDKRVESKYGPPCPAPAEFLGSPS- 119
QY 228 KSEEINEKDLRKKSSELOQTALGNLKOIYYNNSKAITSSSEKSADQFLTNTLLFKGFTTGH 287
Db 120 -----VFLPPPKP-----KDTLMI----- 133
QY 288 WYNDLLDLGSTAATSEYEGSSVDLYGAYGYQCAGGTNKTACMYGGVTLHDNNRLTEE 347
Db 134 -----SRTPEVT-CVVVDVS-----QED 150
QY 348 KKPVINLWIDKQTTVPIDKVKTSKEV-----TVQELDLQARHYLHGKFLYNSDSF 400
Db 151 PEQVFNWYDVG---VEVHNATKPREQFNSTYRVSVVLTVLHQDWLNGK----- 197
QY 401 GGVQVQRLIVFHSSEGSTVSYDLFDAQGOYPTDLLRIYRDNTTISSTLSISLYLTTSI 460
Db 198 -----EY----- 199
QY 461 VMTQPTSLIVSAGDRVTITCKASQSVNDVAWYQOKPGQSPKLLISYTSRYAGVPDRF 520
Db 200 -----KCKVSNK-----GLP--- 209
QY 521 SGGYGTDFTLTISVQAEAAVYFCQDYNPPTFCGGTCKLEIKRADAAPTYSIFPPSS 580
Db 210 -----SSIEK-----TISKAKQPREPQVYTLPPSQ 235
QY 581 EOLTSGGASVWFLNFPKIDINVKKIDGSEKQN-----GVLNSWTDQSDKSDSYSMSS 635
Db 236 EEMTKNOVSLTCLVKGYFSPDIAVENESGQPNENYKTTTPPVLDSD-----DGSFELY 288
QY 636 TLTTLTKDEYERHNSYTCEATHK 657
Db 289 RLTVDKSRWQEGNVFSCSYNME 310

RESULT 44
KV4A_HUMAN STANDARD; PRT; 114 AA.
ID KV4A_HUMAN
AC P01625;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-IV region Len.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=76004342; PubMed=50995;
RA Schneider M., Hilschmann N.;
RT "The primary structure of a monoclonic immunoglobulin-L-chain of
RT subgroup IV of the kappa type (Bence-Jones protein Len).";
RL Hoppe-Seyler's Z. Physiol. Chem. 356:507-557(1975).
RN [2]
RP REVISION TO 9.
RA Salomon A.;
```

RL Submitted (AUG-1996) to the SWISS-PROT data bank.
CC -!- MISCELLANEOUS: THE C REGION OF THIS CHAIN HAS THE INV (3) MARKER.
CC -!- MISCELLANEOUS: THIS IS A BENCE-JONES PROTEIN.
DR PIR: A01903; K4HULN.
DR HSP: P01607; IREI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig_1.
DR SMART: SM00406; Igv: 1.
KW Immunoglobulin V region; Bence-Jones protein.
FT DOMAIN 1 23
FT DOMAIN 24 40
FT DOMAIN 41 55
FT DOMAIN 56 62
FT DOMAIN 63 94
FT DOMAIN 95 101
FT DOMAIN 102 113
FT DISULFID 23 94
FT NON_TER 114 114
SQ SEQUENCE 114 AA; 12640 MW; 0647FD17F236485 CRC64;

Query Match 11.2%; Score 395; DB 1; Length 114;
Best Local Similarity 69.9%; Pred. No. 3.4e-18;
Matches 79; Conservative 12; Mismatches 16; Indels 6; Gaps 1;

QY 460 IVNTQPTSLLSVAGDRVTITCKASQSV-----SNDVAYQKPGQSPKLLISYTSRY 513
DB 2 IVNTQPSDSLAVSLGERATINCKSSQSVLYSSNKNLYAWYQKPGQPKLLIYNASTRE 61

QY 514 AGVPDRFGSGGYGDTFTLTISVQAEADAAYFCQDYNPPTFGGQTKLEIKR 566
DB 62 SGVPDRFGSGSGTDTFTLTISVQAEADAAYVYQVYQVYSPYSGGQTKLEIKR 114

RESULT 45
KV5B_MOUSE
ID KV5B_MOUSE STANDARD; PRT; 136 AA.
AC P01634;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-V region MOPC 21 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=82059477; PubMed=6170937;
RA Hamlyn P.H., Gait M.J., Milstein C.;
RT "Complete sequence of an immunoglobulin mRNA using specific priming
RT and the dideoxynucleotide method of RNA sequencing.";
RL Nucleic Acids Res. 9:4485-4494(1981).
RN [2]
RP SEQUENCE OF 30-136.
RX MEDLINE=73053310; PubMed=4638343;
RA Svasti J., Milstein C.;
RT "The complete amino acid sequence of a mouse kappa light chain.";
RL Biochem. J. 128:427-444(1972).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: V00810; CAA24192.1; ALT_TERM.
DR PIR: A01917; KVM21.
DR HSP: P80362; IWL.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.

DR Pfam: PF00047; Ig: 1.
DR SMART: SM00406; Igv: 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 29
FT CHAIN 30 136
FT DOMAIN 30 52
FT DOMAIN 53 63
FT DOMAIN 64 78
FT DOMAIN 79 85
FT DOMAIN 86 117
FT DOMAIN 118 126
FT DOMAIN 127 136
FT NON_TER 136 136
SQ SEQUENCE 136 AA; 14902 MW; 8CDD85113996D1C2 CRC64;

Query Match 11.2%; Score 395; DB 1; Length 136;
Best Local Similarity 60.9%; Pred. No. 4.3e-18;
Matches 81; Conservative 18; Mismatches 34; Indels 0; Gaps 0;

QY 433 TLLRIYRDNTTSSLSISLYTTSIVMTQPTSLLSVAGDRVTITCKASQSVNDVA 492
DB 4 TSMGKIMESHVTLVFISILCLYGACGNIVMTQSPKMSVSGERVTLTCKASENVVTYVS 63

QY 493 WYQKPGQSPKLLISYTSRYAGVPDRFGSGGYGDTFTLTISVQAEADAAYFCQDYN 552
DB 64 WYQKPEQSPKLLIYGASNRVTGVPDRFTGSGSATDFTLTISVQAEADLYHCGGY 123

QY 553 PTFGGGQTKLEIK 565
DB 124 PYTFGGGQTKLEIK 136

RESULT 46
HV52_MOUSE
ID HV52_MOUSE STANDARD; PRT; 117 AA.
AC P06327;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig heavy chain V region VH558 A1/A4 precursor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85099340; PubMed=2578321;
RA Yancopoulos G.D., Alt F.W.;
RT "Developmentally controlled and tissue-specific expression of
RT unrearranged VH gene segments.";
RL Cell 40:271-281(1985).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M13787; AAA38499.1; -.
DR PIR: A02029; HVMSA1.
DR HSP: P01810; 2FBI.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003596; Ig_V.
DR Pfam: PF00047; Ig: 1.
DR SMART: SM00406; Igv: 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 117
FT DOMAIN 20 49
FT DOMAIN 50 54
FT DOMAIN 55 68

```
FT DOMAIN 69 85 COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 86 117 FRAMEWORK-3.
FT DISULFID 41 115 BY SIMILARITY.
FT NON_TER 117 117
SQ SEQUENCE 117 AA; 12971 MW; 8B0BC138856DFC9D CRC64;

Query Match 11.1%; Score 392; DB 1; Length 117;
Best Local Similarity 75.5%; Pred. No. 5.4e-18;
Matches 74; Conservative 12; Mismatches 12; Indels 0; Gaps 0;

Qy 1 EVOLQSGPDLVKPCASVKISKASGYSTGYMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 QVQLQSGPELVKGLKPKALVKISKASGYFTSYDINWVKRPGQGLEWIGWYPGDGSTKY 79

Qy 61 NOKFKDKATLVKDSSTTAYMELRSLTSEDSAVVYCARG 98
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 NEKFKGKATLVKDSSTTAYMELRSLTSEDSAVVYCARG 117

RESULT 47
HV1C_HUMAN STANDARD; PRT; 147 AA.
AC P01744;
DT 21-JUL-1986 (Rel. 01, Created)
DE 16-OCT-2001 (Rel. 40, Last sequence update)
DE Ig heavy chain V-I region ND precursor (Fragments).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83065234; PubMed=6815656;
RA Kenten J.H., Molgaard H.V., Houghton M., Derbyshire R.B., Viney J.,
RA Bell L.O., Gould H.J.;
RT "Cloning and sequence determination of the gene for the human
RT immunoglobulin epsilon chain expressed in a myeloma cell line.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:6661-6665(1982).
RN [2]
RP SEQUENCE OF 20-147.
RA Bennich H.H., Johansson S.G.O., von Bahr-Lindstrom H.;
RL (In) Bach M.K. (eds.);
RL Immediate hypersensitivity: modern concepts and developments, pp.1-36,
RL Marcel Dekker, New York (1978).
CC -I- MISCELLANEOUS: THIS EPSILON CHAIN WAS ISOLATED FROM A MYELOMA
CC PROTEIN.
DR PIR; A02026; ELHUND.
DR HSSP; P01789; IMCP.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region; Signal.
FT SIGNAL 1 19
FT CHAIN 20 147 IG HEAVY CHAIN V-I REGION ND.
FT MOD_RES 20 20 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 41 115
FT CONFLICT 21 21 T -> V (IN REF. 2).
FT CONFLICT 53 54 IH -> HI (IN REF. 2).
FT CONFLICT 67 68 VG -> GV (IN REF. 2).
FT CONFLICT 125 125 MISSING (IN REF. 2).
FT NON_TER 147 147
SQ SEQUENCE 147 AA; 16491 MW; 948F9F72A5366C20 CRC64;

Query Match 10.9%; Score 385; DB 1; Length 147;
Best Local Similarity 55.5%; Pred. No. 2e-17;
Matches 71; Conservative 24; Mismatches 25; Indels 8; Gaps 1;

Qy 1 EVOLQSGPDLVKPCASVKISKASGYSTGYMHWKQSPCKGLEWIGRINPNNGVTLY 60
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
20 QTQLVQSGAEVRKPGASRVSKASGYFTSYDTHWIRQAPGHGLEWYGNPNSSGNTY 79
```

```
Qy 61 NOKFKDKATLVKDSSTTAYMELRSLTSEDSAVVYCARG 112
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
80 APFQGRVMTTRDASFSTAYMDLSRLSDSDSAVFYCAKSDPFWSDYNFYSYTLDDVWGO 139

Qy 113 GTSVTVSS 120
Db :|||||
140 GTTVTVSS 147

RESULT 48
GC2_HUMAN STANDARD; PRT; 326 AA.
ID GC2_HUMAN
AC P01859;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-2 chain C region.
DE IGHG2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 2-326 FROM N.A.
RX MEDLINE=82197621; PubMed=6804948;
RA Ellison J.W., Hood L.E.;
RT "Linkage and sequence homology of two human immunoglobulin gamma
RT heavy chain constant region genes.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:1984-1988(1982).
RN [2]
RP SEQUENCE OF 88-115 FROM N.A.
RC TISSUE=Fetal liver;
RX MEDLINE=83001943; PubMed=6811139;
RA Takahashi N., Ueda S., Obata M., Nikaido T., Nakai S., Honjo T.;
RT "Structure of human immunoglobulin gamma genes: Implications for
RT evolution of a gene family.";
RL Cell 29:671-679(1982).
RN [3]
RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.
RC TISSUE=Fetal liver;
RX MEDLINE=84235992; PubMed=6329676;
RA Krawinkel U., Rabbitts T.H.;
RT "Comparison of the hinge-coding segments in human immunoglobulin gamma
RT heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass
RT genes.";
RL EMBO J. 1:403-407(1982).
RN [4]
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
RX MEDLINE=81007873; PubMed=6774012;
RA Wang A.-C., Tung E., Fudenberg H.H.;
RT "The primary structure of a human IgG2 heavy chain: genetic,
RT evolutionary, and functional implications.";
RL J. Immunol. 125:1048-1054(1980).
RN [5]
RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIE).
RX MEDLINE=80001357; PubMed=113060;
RA Connell G.E., Parr D.M., Hofmann T.;
RT "The amino acid sequences of the three heavy chain constant region
RT domains of a human IgG2 myeloma protein.";
RL Can. J. Biochem. 57:758-767(1979).
RN [6]
RP SEQUENCE OF 238-275 (ZIE).
RX MEDLINE=80114419; PubMed=118920;
RA Hofmann T., Parr D.M.;
RT "A note of the amino acid sequence of residues 381-391 of human
RT immunoglobulin gamma chains.";
RL Mol. Immunol. 16:923-925(1979).
RN [7]
RP REVISIONS TO 25: 59; 60 AND 264-268 (ZIE).
RA Hofmann T., Parr D.M.;
RL Submitted (MAR-1980) to the PIR data bank.
RN [8]
RP SEQUENCE OF 1-121 (DOT).
```

```

RX MEDLINE=9525298; PubMed=7737190;
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT "Characterization of the two unique human anti-flavin monooxonal
RL immunoglobulins.";
RN Eur. J. Biochem. 228:886-893(1995).
RP [9]
RP DISULFIDE BONDS.
RX MEDLINE=7203350; PubMed=4940472;
RA Milstein C., Frangione B.;
RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";
RL Biochem. J. 121:217-225(1971).
RN [10]
RP DISULFIDE BONDS.
RX MEDLINE=69064124; PubMed=5782707;
RA Frangione B., Milstein C., Pink J.R.L.;
RT "Structural studies of immunoglobulin G.";
RL Nature 221:145-148(1969).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J00230; AAB59393.1; -.
DR PIR; A02148; G2HU.
DR HSSP; P01857; 1FC1.
DR GENE; HGNC:5526; IGHG2.
DR MIN; 147110; -.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00410; Ig_like; 1.
DR SMART; SM00407; IGH1; 2.
DR PROSITE; PS00290; IG_MHC; 2.
DR Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DOMAIN 1 98
FT CH1.
FT HINGE.
FT DOMAIN 99 110
FT DOMAIN 111 219
FT CH2.
FT DOMAIN 220 326
FT CH3.
FT DISULFID 14 14
FT INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 83
FT INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 102 102
FT INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 103 103
FT INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 106 106
FT INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109
FT INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 140 200
FT DISULFID 246 304
FT DISULFID 246 304
FT SITE 156 156
FT MOD_RES 326 326
FT VARIANT 60 60
FT CONFLICT 109 109
FT SEQUENCE 326 AA; 35884 MW; 8310878C6878CF9C CRC64;
Query Match 10.7%; Score 378; DB 1; Length 326;
Best Local Similarity 22.2%; Pred. No. 1.6e-16;
Matches 123; Conservative 56; Mismatches 113; Indels 262; Gaps 14;
QY 121 AKTTPSVYPLAPGSAATNSVTLGCLVKGVEPPTVTWNSGSLSSGVHTFPVAVLQSD 180
DB 1 ASTKGPSVFPLAPCSRSTSESTAALGCLVKDYFPEPTVSWNSGALTSGVHTFPAVLQSS 60
QY 181 -LYTSSSVTPSPSTWPTSETVCNVAHPASSTKVDKIVPR-----DSGGPSEK 228
DB 61 GLYSLSVTVTPSSNFGQTYTCNVNDRKPSNTKVDKTVKRCVPCPAPPVAGPS-- 118
QY 229 SEETNEKDLRKSLQGTALGNLKOIYYNSKATTSSEKSADQFTLNTLLKGFPTGHPW 288

```

```

Db 119 -----VFLFPPKP-----KDTLMI----- 132
QY 289 YNDLLVLDLGGTAATSEYEGSSVDLYGAYGYQCAGGTENKTCACMYGVTLLHNNRLTEEK 348
DB 133 -----SRTFEVT-CVVVDVS-----HEDP 150
QY 349 KVPINLWIDGKQTTVPIDKVKTSKEVTQVQLDLQARHYLHGKFGLYNSDSFGGKVQRL 408
DB 151 EVQFNWYVDG---VEVHNKTKPRE----- 172
QY 409 IVFHSSEGSTVSYDLFDAQQYPTDLLRIYRDNTTISSTLSISLYLYTTSIVMTQTPTS 468
DB 173 -----EQFNSTFRVSVLTVVHQD----- 191
QY 469 LLVSAGDRVTITCRASQSVSDNVAVYQOKPGOSPKLLISYTSRRYAGVYDREFSGSGYGT 528
DB 192 -----WLNKEYKC-----KVSNNKGLPAP 210
QY 529 FTLTSSVQAEDAAYVFCQDYNSPPTFGGKTKLEIKRADAAPTYSIFPPSSEQLTSGGA 588
DB 211 IKTISKTKGQ-----PREPQVYTLPPSREEMTKNQV 242
QY 589 SVVCFLNFPKIDINVKKIDGSEKQ-----GVLSNWTDDQSKDSTYSMSSTLTLTDE 643
DB 243 SLTCLVKGFPDIAVEWESNGQPNENYKTTTPMLDS-----DGSFLYSKLTVDKSR 295
QY 644 YERNHNSYTCATHK 657
DB 296 WQGNVFCSCVWHE 309
RESULT 49
KVLM_HUMAN
ID KVLM_HUMAN STANDARD; PRT; 108 AA.
AC P01605;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Ig kappa chain V-I region Lay.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE.
RX MEDLINE=77038198; PubMed=824717;
RA Capra J.D., Klapper D.G.;
RT "Complete amino acid sequence of the variable domains of two human
RT IgM anti-gamma globulins (Lay/pom) with shared idiotypic
RT specificities.";
RL Scand. J. Immunol. 5:677-684(1976).
CC -I- MISCELLANEOUS: THE SECOND AND THIRD HYPERVARIABLE REGIONS OF THIS
CC CHAIN ARE IDENTICAL WITH THOSE OF THE HUMAN POM V-III KAPPA CHAIN,
CC WITH WHICH IT SHARES CERTAIN IDIOTYPIC DETERMINANTS.
CC -I- MISCELLANEOUS: THIS CHAIN WAS ISOLATED FROM AN IGM WITH ANTI-GAMMA
CC GLOBULIN ACTIVITY.
DR PIR; A01871; KLHULY.
DR HSSP; P01607; 1REI.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00406; IGV; 1.
KW Immunoglobulin V region.
FT DOMAIN 1 23
FT DOMAIN 24 34
FT COMPLEMENTARITY-DETERMINING-1.
FT DOMAIN 35 49
FT COMPLEMENTARITY-DETERMINING-2.
FT DOMAIN 50 56
FT COMPLEMENTARITY-DETERMINING-3.
FT DOMAIN 57 88
FT COMPLEMENTARITY-DETERMINING-4.
FT DOMAIN 89 97
FT COMPLEMENTARITY-DETERMINING-5.
FT DISULFID 98 107
FT DISULFID 23 88
FT NON_TER 108 108
SQ SEQUENCE 108 AA; 11834 MW; 739993A95431434A CRC64;

```


QY 517 PDRFSGSGYGTDFTLTISVQAEDAAYVFCQDYNSP-TEGGGTTKLEIKR 566

THIS PAGE BLANK (USPTO)

Result No.	Query			ID	Description
	Score	Match	Length		
1	1010	28.7	463	Q991C4	Q991c4 mus musculus
2	948	26.9	260	Q931M4	Q931m4 staphylococcus aureus
3	929.5	26.4	238	11 Q8VC16	Q8vc16 mus musculus
4	924.5	26.2	473	11 Q9DB14	Q9db14 mus musculus
5	918	26.1	260	16 Q99S03	Q99s03 staphylococcus aureus
6	917.5	26.1	473	11 Q99125	Q99125 mus musculus
7	910	25.8	239	11 Q8VC55	Q8vc55 mus musculus
8	909.5	25.8	238	11 Q99M37	Q99m37 mus musculus
9	908	25.8	234	11 Q91WF8	Q91wf8 mus musculus
10	902	25.6	468	11 Q991J31	Q991j31 mus musculus
11	900	25.6	233	11 Q91W89	Q91w89 mus musculus
12	896	25.4	234	11 Q8R062	Q8r062 mus musculus
13	893	25.4	278	11 Q921K1	Q921k1 mus musculus
14	889	25.2	469	11 Q8R3V9	Q8r3v9 mus musculus
15	888.5	25.2	474	11 Q8R3H6	Q8r3h6 mus musculus
16	881.5	25.0	235	11 Q91W12	Q91w12 mus musculus

```
Db 20 QVOLOQSGAELARPAGSARVLSCKASGYTFTGYGVSWKQRTQGGLWEHGEIYPGSGNTYY 79
Qy 61 NOKFKDKATLTVDKSTAYMELRSLTSEDSAVYICARSTMTITNVMDYWGQGTSVTVSS 120
Db 80 SEKFKGKATLTITDKSSSTAYMHLSELTSSEDSAVYICARSSYYSDLFAYWGQGTTLTVSA 139
Qy 121 AKTTPPSVYPLAPGSAQAQTNMVTGLGKVGKGYFPEPVTVTNWNSGLSSGVHTFPFAVLQSD 180
Db 140 AKTTPPSVYPLAPGSAQAQTNMVTGLGKVGKGYFPEPVTVTNWNSGLSSGVHTFPFAVLQSD 199
Qy 181 LYTLLSSVTPSPSTWPTSETVTCNVAHPASSTKVDKKIIVPRDSGGPSEKSEINEKDLRKK 240
Db 200 LYTLLSSVTPSPSTWPTSETVTCNVAHPASSTKVDKKIIVPRDCG----- 242
Qy 241 SELOQTALGNLKOIYYNSKAITSEKSDAQFLTNTLFLKGFETGHPWYNLLVDLGSTA 300
Db 243 CKPCICTVPEVSSVFIFFPK-----PKVLTITL----- 271
Qy 301 ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKKVPINLWIDGKQ 360
Db 272 -----TP-KVTCVVVDIS-----KDDPEVQFSWFVD--- 296
Qy 361 TTVPIDKVKTSKEV-----TVQELDLQARHYLHGK-FGL-YNSDSFGGKVQRLIVF 411
Db 297 -DVEYHTAQTPREPQFNSTFRSVSELPIHMDHDLNGLKREKFCRVNSAAPPAPIER----- 350
Qy 412 HSSEGSTVSYDLFDAQGOVDPDLLRIYRDNTTISSTLSLSISLYLTTISVMTQTPTSLLV 471
Db 351 -----TISKT----- 355
Qy 472 SAGDRVTITCKASQSVSNDAVYQKQPGSKLLISYTSRSGYAGVPDRFSGSGYGTDFTL 531
Db 356 ----- 355
Qy 532 TISSVOAEDAAYVFCQDYNSPPTGGGKLEIKRADAAPTVIFPPSPSEQLTSGGASVW 591
Db 356 -----KGRKAPQVYTIPTPPKQMAKDVKSLT 382
Qy 592 CFLNFPYKPDINVKKIDGSRQNGVLNSWTDDQSKDSTYSMSSTLTLTCKDEYERHNSVT 651
Db 383 CMITDFPEDITVEQWNGQPAEN-YKNTQPMIDT-DGSYFIYSLNOKSVNWEAGNTFT 440
Qy 652 CEATHK 657
Db 441 CSVLHE 446
RESULT 2
Q931M4 ID Q931M4 PRELIMINARY; PRT; 260 AA.
AC Q931M4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Enterotoxin P.
GN SEP OR SAV1948.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hoshiyama A.,
RA Mizutani-U. Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RL aureus.";
RL Lancet 357:1225-1240(2001).
```

```
DR EMBL; AP003364; BAB58110.1; -.
DR InterPro; IPR001961; Strep_Strep_toxin.
DR Pfam; PF01123; Strep_Strep_toxin; 1.
DR Pfam; PF02876; Strep_Strep_toxin; 1.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 260 AA; 30016 MW; 15C2D36270FA8241 CRC64;
Query Match 26.9%; Score 948; DB 16; Length 260;
Best Local Similarity 76.4%; Pred. No. 2.2e-51;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;
Qy 226 SEKSEINEKDLRKKSELOQTALGNLKOIYYNSKAITSEKSDAQFLTNTLFLKGFETG 285
Db 28 SEKSEINEKDLRKKSELOQTALGNLKOIYYNEKAKTENKESHQFLQHTILFKGFTD 87
Qy 286 HPWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 345
Db 88 HSWYNLLVDLFDSDKDIVDKYKGVLDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 147
Qy 346 BEKKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQ 405
Db 148 BEKKVPINLWIDGKQNTVPLETVTNKNKNTVQELDLQARRYLQEKYLNLYNSDVFQGVQ 207
Qy 406 RGLIVFHSSEGSTVSVDLFDAGGOVDPDLLRIYRDNTTISSTLSLSISLYLYTT 458
Db 208 RGLIVFHSSTEPSVNDLFGAQQGYSNTLLRIYRDNTINSENMHIDIYLYTS 260
RESULT 3
Q8VC16 ID Q8VC16 PRELIMINARY; PRT; 238 AA.
AC Q8VC16;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 26.2 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC019760; AAH19760.1; -.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003597; IG_C1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; IG_V.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGc1; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 238 AA; 26224 MW; 35EC08E3DE5414AD CRC64;
Query Match 26.4%; Score 929.5; DB 11; Length 238;
Best Local Similarity 81.8%; Pred. No. 2.7e-50;
Matches 180; Conservative 15; Mismatches 20; Indels 5; Gaps 2;
Qy 457 TTSIVMTQTPPSLVASAGDRVTITCKASQSV--SND---VAVYQKQPGSKLLISYTS 511
Db 18 SSDVYMTQTPSLPSVLDGQASISCRSSQSLVHNSGNTYLLHWYLOKQPGSKLLIYKVN 77
Qy 512 RYAGVPDRFSGSGYGTDFTLTISVQAEADAAVYFCQDYNSPPTFGGKLEIKRADAA 571
Db 78 RFGVPDRFSGSGGTDFTLTKISRVEADLGVYFCSQSTHVPPTFGGKLEIKRADAA 137
Qy 572 TVSIFPPSSEQLTSGGASVVCFLNNFYPKIDGSRQNGVLNSWTDDQSKDSTY 631
```

```

Db 138 TVSIIPPSSEQLTSGASVYCFVNNFYKPDINVKWKIDGSRQNGVLSWTDQDSDKSTY 197
QY 632 SMSSTLTTLTKDEYERHNSVTCETHKTSPIVKSFNRE 671
Db 198 SMSSTLTTLTKDEYERHNSVTCETHKTSPIVKSFNRE 237

RESULT 4
Q9D8L4
ID Q9D8L4 PRELIMINARY; PRT; 473 AA.
AC Q9D8L4
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE 181006009Rik protein.
GN IGH-1 OR 181006009RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Arawaka T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007918; BAB25349.1; -.
DR HSSP; P01842; 7FAB.
DR MGD; MGI:96443; Igh-1.
DR InterPro; IPR003599; Igh.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_Y.
DR Pfam; PF00047; Igh_4.
DR SMART; SM00409; IgC; 2.
DR SMART; SM00407; IGc1; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_Like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 51699 MW; 9DED57A514475FB8 CRC64;

Query Match 26.2%; Score 924.5; DB 11; Length 473;
Best Local Similarity 33.8%; Pred. No. 1.4e-49;
Matches 239; Conservative 53; Mismatches 123; Indels 293; Gaps 15;

QY 1 EVQLQSGDPLVKPGASVKISCKASGYSTGYIMHWKSPGKGLIEWIGRINPNNGVTLY 60
Db 20 QVQLKQSGAEVLPKPGASVKISCKASGYSTGYIMHWKSPGKGLIEWIGRINPNNGVTLY 79
QY 61 NQKFKDKATLTVDKSTTAYMELRSITSEDSAVYVCARSTMTITNYMD---YWGQGTSTV 117
Db 80 NEKFKGKATLTADKSSSTAYMQLSSLTSEDSAVYFCARS-----GYDYDMFWAYWGQGT 135

```

```

QY 118 VSSAKTTTPSVVYPLAPGSAQAQNSMVTLCGLVKGYFPEPVTVTNWNSGSLSSGVHTFPAVL 177
Db 136 VSAAKTTAPSVVYPLAPGCGTGTSSVTLGCLVKGYFPEPVTVTNWNSGSLSSGVHTFPAVL 195
QY 178 QSDLYTLSSSVTVPSSTWPSSETVTCNVAHPASSTKVDKIVPRDGGGPGSEKSEENKRD 237
Db 196 QSGLYTLSSSVTVTNTWPSQTITCNVAHPASSTKVDKIEPR----- 238
QY 238 RKKSELQGTALGNLKOIYYVNSKAITSSEKSDAQFLTNLLFKGFFTHGHPWYNLDLVDLG 297
Db 239 ----- 238
QY 298 STAATSEYEGSVLDYGAYGYQCAGGTPNKTCMYGGVTLHDNNRLTEBKVPINLWID 357
Db 239 -----VPI----- 241
QY 358 GKQTTVPIDKVKTSKKEVTVOELDQARHYLHGKFGLYNSDSFGKVGKQGLIVFHSSEGS 417
Db 242 ---TQNPCPLKECPCAA-----PDLGGP----- 264
QY 418 TVSYDLEDAQGOVPTDLLRIYRDNTTISSTLSISLYLVTTSIVMTQTPTSLVSAGDRV 477
Db 265 ----SVFIFPPKIKDLM-----ISLSPWTCVV----- 290
QY 478 TITCKASQSVNDVAMVQKPGQSPKLLISYTSRYAGVDFRFGSGYGTDTLTISVYQ 537
Db 291 -----DPDQISW-----FVNNVE 309
QY 538 AEDAAVYFCQDYNSPPTFGGKTKLEIKRAD----- 568
Db 310 VHTAQTTQTHREDYNS--TLRVVSALPIQHDWMSGKEFKCKVNNRALPSPIEKTIKPRG 367
QY 569 --AAPTSTVSIFFSSQLTSGGASVVCFLNFKYKIDINVKWKIDGSRQNGVLSWTDQDS 626
Db 368 PYRAQVYVLPPEAEMTKKESLTCMIMGFLPAELAVDWTNGRTEQN-YKNTATVLDS 426
QY 627 KDTYSMSSTLTTLTKDEYERHNSVTCETHK-----TSTSPVKSFN 669
Db 427 -DGSYFMYSKLRVQKSWERGSFLFACSVVHEGLNHLTKTISRSLGK 473

RESULT 5
Q99SU3
ID Q99SU3 PRELIMINARY; PRT; 260 AA.
AC Q99SU3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Enterotoxin P.
GN SEP OR SAI761.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RC SEQUENCE FROM N.A.
RX MEDLINE=213111952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hosoyama A., Mizutani-U I Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003135; BAB43036.1; -.
DR HSSP; P13163; 1SYT.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Stp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.

```

DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Complete proteome.
SQ SEQUENCE 260 AA; 29708 MW; 087C5B4EC028CFDB CRC64;

Query Match 26.1%; Score 918; DB 16; Length 260;
Best Local Similarity 73.0%; Pred. No. 1.6e-49;
Matches 170; Conservative 21; Mismatches 42; Indels 0; Gaps 0;

QY 226 SEKSEINEKDLRKSELOGTALGNLKIYYIYNSKAITSSSEKSAOFLTNTLLFKGFTG 285
DB 28 SEKSEINEKDLRKSELOGTALGNLKIYYIYNSKAITSSSEKSAOFLTNTLLFKGFTG 87
QY 286 HPWYNDLLVDLSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 345
DB 88 HPWYNDLLVDLSTAKDTANIYKGGKVDLYGAYGYOCTGGTPFKTACMYGGVTLHDNNQLE 147
QY 346 EEKKVPINLWIDGKQTPIDPKVYKSKKEVTVOELDLQARHYLHGKFLYNSDSFGKQV 405
DB 148 EEKKVPINLWIDGKQTPVPLGTNKKKEVTVOELDLQSRHYLHETYNLYNTDAENGKIQ 207
QY 406 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRONTTISSTLSISLYLYTT 458
DB 208 RGLIEFHPSSGDSVGYDLFGAQQYPTQLRIYRONTTIKSKNMHIDIYLYTT 260

RESULT 6
Q99L25 PRELIMINARY; PRT; 473 AA.
AC Q99L25;
DT 01-JUN-2001 (TremBLrel. 17, Created)
DT 01-JUN-2001 (TremBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Similar to RIKEN CDNA 181060009 gene.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003888; AAH03888.1; -.
DR HSSP; P01842; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003600; Ig-like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; IGC1; 3.
DR SMART; SM00406; IGV; 1.
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
SQ SEQUENCE 473 AA; 52449 MW; BE9889B79860A155 CRC64;

Query Match 26.1%; Score 917.5; DB 11; Length 473;
Best Local Similarity 33.8%; Pred. No. 3.8e-49;
Matches 232; Conservative 59; Mismatches 141; Indels 255; Gaps 12;

QY 1 EYVLOQSGDLVKPGASKVSKSGYFTGYIMHWKQSPGKGLWIGRINPNNGVTLY 60
DB 20 QYVLOQSDAELVKPGASKVSKSGYFTDTHIHWKQSPGKGLWIGYIPRGSTKY 79
QY 61 NQKFKDKATLVKDSSTTAYMELRSLTSEDSAVYYCARSTMTN---YVMDYWGQGTSV 116
DB 80 NEKFKGKATLTADKSSATYMLNSLTSEDSAVCFCSRGSGIYGYLYFDYWGQGTI 139
QY 117 TVSSAKTTPPSVYPLAPGSAAGTNSWVLGCLVKGYPFEPVTVTWNSSLSGVTFFAV 176
DB 140 TVSSAKTTPPSVYPLAPGCGDGTGSSVTLGCLVKGYPFEPVTVTWNSSLSGVTFFAV 199

QY 177 LOSDLTYLTSSSVTPSPSTWSETVCNVAHPASSTKVDKKIVPRDSGGSPSEKSEBINEKD 236
DB 200 LOSDLTYLTSSSVTPSPSTWSETVCNVAHPASSTKVDKKIEPR---GPTIKPCP----- 251
QY 237 LRKKELOGTALGNLKIYYIYNSKAITSSSEKSAOFLTNTLLFKGFTGHPWYNDLLVDL 296
DB 252 ----- 251
QY 297 GSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWI 356
DB 252 -----PCKCPAPN-----LLGGPSVF----- 267
QY 357 DGKQTPIDPKVYKSKKEVTVOELDLQARHYLHGKFLYNSDSFGKQVORGLIVFHSSEG 416
DB 268 -----IFPPKIKDVLM-----SLS 282
QY 417 STVSVDLFDAGQYPTDLLRIYRONTTISSTLSISLYLYTTISYVMTQPTPTSLVSAGDR 476
DB 283 PMVTCVWVDVSEDDPD-----VQISWFVNNVEVLTAQTQT----- 317
QY 477 VTITCKASQSVSNDAVYQKQSPKLLISYTSRYAGVPRDFSGSGYGTDTLTITISV 536
DB 318 -----HREDYNSTLRVV 329
QY 537 QA-----EAAVYFCQDYNSPTFTGGTKLEIKRADAAPTVSIPTSPSEQLTSGA 588
DB 330 SALPTQHQDWMKSGKEFKCKVNNKALPAPIERTISKPKSVRAPOVYVLPPEEMTKKQV 389
QY 589 SVVCFNNFYPKDINVKWIDGSERON-----GVLSNWTDDQSKDSTYSMSSTLTLTDE 643
DB 390 TLTCTMVTDFMPEDIYVEVNTNGKTELNYKNTPEVLDS-----DCSYFMYSKLRVEKN 442
QY 644 YERHNSYTCETHK--TSTSPIVKSNR 669
DB 443 WVERNSYSCVVVHEGLHNHHTKSF 469

RESULT 7
Q8VC55 PRELIMINARY; PRT; 239 AA.
AC Q8VC55;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hypothetical 26.3 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC021781; AAH21781.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; IGC1; 1.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26303 MW; C16119CACA25C337 CRC64;

Query Match 25.8%; Score 910; DB 11; Length 239;
Best Local Similarity 76.0%; Pred. No. 4.4e-49;
Matches 177; Conservative 21; Mismatches 25; Indels 10; Gaps 2;

QY 449 LSIISLYLT-----TSIVMTQPTPTSLVSAGDRVTITCKASQSVSN-----VAWYQK 498
DB 449 LSIISLYLT-----TSIVMTQPTPTSLVSAGDRVTITCKASQSVSN-----VAWYQK 498

RESULT 12

```

Q8R062      PRELIMINARY;          PRT;      234 AA.
ID  Q8R062
AC  Q8R062;
DT  01-JUN-2002 (TrEMBLrel. 21, Created)
DT  01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT  01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE  Hypothetical 25.9 kDa protein.
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae
OX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=COLON;
RA  Strausberg R.;
RL  Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RR  EMBL; BC027418; AAH27418.1; -.
RW  Hypothetical protein.
KW  SEQUENCE      234 AA; 25857 MW; 4EB08C81426AEAB1 CRC64;
SQ  SEQUENCE      25.4k; Score 896; DB 11; Length 234;
      Best Local Similarity 81.6%; Pred. No. 3.2e-48;
      Matches 173; Conservative 13; Mismatches 26; Indels

```

Db	22	IQMTQTSSLSASLDGRVITSCASOGISNYLNWYQKPDGTGTVKLLIYTTSSLUHGVPSR	81
Qy	520	FSGSGYGTDFTLTSSVOAEDAAVYFCQODYNSPPTFGGKLEIKRAADAAPTVSIFPPS	579
Db	82	FSGSGGTHYSLTISNLEPDIATYCYCQSQPPTFGGKLEIKRAADAAPTVSIFPPS	141
Qy	580	SEQLTSGGASVWVCFLLNFYPKDINVKWKIDGSRQGVLSNWTQDQSKDSTYSMSSTLTL	639
Db	142	SEQLTSGGASVWVCFLLNFYPKDINVKWKIDGSRQGVLSNWTQDQSKDSTYSMSSTLTL	201
Qy	640	TKDEYERHNSYTCEATHKSTSPIVKSFNRNE	671
Db	202	TKDEYERHNSYTCEATHKSTSPIVKSFNRNE	233
RESULT 13			
Q921K1			
ID	Q921K1	PRELIMINARY;	PRT; 278 AA.
AC	Q921K1;		
DT	01-DEC-2001 (TremblRel. 19, Created)		


```
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Unknown (protein for MGC:18977).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC012207; AAH12207.1; -.
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig_2.
SQ SEQUENCE 278 AA; 29778 MW; F894F955DDCD948A CRC64;

Query Match 25.4%; Score 893; DB 11; Length 278;
Best Local Similarity 76.8%; Pred. No. 6.1e-48;
Matches 169; Conservative 18; Mismatches 33; Indels 0; Gaps 0;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYMHVWVKSPGKGLGWIGRINPNNGVTLY 60
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
20 QVLOQPGAEVLKPGASVKLSCKASGYFTSYMHVWVKRRGQGLGWIGNINPNSGNTY 79
QY 61 NOKFKDKATLVTDKSTTAYMELRSLTSEDSAVYYCARSTMITINYYMDYWGQGTSTVSS 120
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
80 NEKFNKATLAVDKSSSTVYMQLSLTSEDSAVYYCTRGVYDDVYDVYMGAGTTVTSS 139
QY 121 AKTTPSVVPLAPGSAQAQNSMTVLCGLVKGYPEPVTYTNWNSGSLSSGVHHPFAVLQSD 180
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
140 AKTAPSVVPLAPVCGGTTGSSVTLGCLVKGYPEPVTYTNWNSGSLSSGVHHPFALLQSG 199
QY 181 LYTLLSSSVTPSPSETVTCNVAHPASSTKVDKKIVPR 220
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
200 LYTLLSSSVTPSPSETVTCNVAHPASSTKVDKKIEPR 239

RESULT 14
Q8R3V9 PRELIMINARY; PRT; 469 AA.
AC Q8R3V9;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 52.0 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024405; AAH24405.1; -.
KW Hypothetical protein.
SQ SEQUENCE 469 AA; 51976 MW; 534793F155D05457 CRC64;

Query Match 25.2%; Score 889; DB 11; Length 469;
Best Local Similarity 32.9%; Pred. No. 2.2e-47;
Matches 221; Conservative 68; Mismatches 129; Indels 254; Gaps 16;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYMHVWVKSPGKGLGWIGRI-NPNNGVTLY 59
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
20 EVNLVESGGVLQPGGSLRLSCAASGFTFDYYSVNVPPGPKALEWLGFIKRNKANGYTT 79
QY 60 -NOKFKDKATLVTDKSTTAYMELRSLTSEDSAVYYCARSTMITINYYMDYWGQGT 114
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
80 EYSASVKGRTTISRDNSQSLYLQMLNRAEDSATIYCARRDRSSYYISGTSFAYWGQGT 139
QY 115 SVTVSSAKTTPSVVPLAPGSAQAQNSMTVLCGLVKGYPEPVTYTNWNSGSLSSGVHHP 174
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
140 LVTVAATTPSVVPLAPGSAQAQNSMTVLCGLVKGYPEPVTYTNWNSGSLSSGVHHP 199
```

```
QY 175 AVLQSDLYTLSSSVTPSPSETVTCNVAHPASSTKVDKKIVPRDSGGSPSEKSEINE 234
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
200 AVLQSDLYTLSSSVTPSPSETVTCNVAHPASSTKVDKKIVPRDCG----- 248
QY 235 KDLRKSELOQTALGNLKOIYYNSKAITSSSEKSAQDQFLTNLLPKGFTTGHWPYNDLLV 294
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
249 -----CKPCICTVPEVSSVFIFPPK-----PKDVLITL----- 277
QY 295 DLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTKACMYGGVTLHDNRLTEKKVPINL 354
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
278 -----TP-KVTCVVVDIS-----KDDPEVQFQSW 299
QY 355 WIDGKOTTPIDKVTSSKEV-----TVQELDLQARHYLHGK-EGL-YNDSDFGKQVQ 405
DB :| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
300 FVD---DVEVHTAQTTPREEQFNSTERSVSELPIMHQDWLNGKEFKCRVNSAAFPAPIE 355
QY 406 RGLVHPSSEGSSTVSYDLFAQGOYPTDLLRIYRDNTTISSTLSLSLYTTSIVMTQT 465
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
356 K-----TISK----- 361
QY 466 PISLLVSAGDRVITTCASQSVSNVAVYQKPGQSPKLLISVTSRYAGVPDRFSGSGY 525
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
362 -----KGRPKAPQVVTIPPPKEQMAK 382
QY 586 GGASVVCFLNFPKIDNVKWKIDGSEKONGVLSNWDQSDKSTYSMSSTLTLTDEYE 645
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
383 DKVSLTCMTIDFPEDITVEWQWNGQPAEN-YKNTQPIQMDT-DGSYFVYSKLVQKSNWE 440
QY 646 RHNSYTCETHAK 657
DB : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
441 AGNTFTCSVLHE 452

RESULT 15
Q8R3H6 PRELIMINARY; PRT; 474 AA.
AC Q8R3H6;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Hypothetical 51.7 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025447; AAH25447.1; -.
KW Hypothetical protein.
SQ SEQUENCE 474 AA; 51748 MW; 8608B57C6CD2874A CRC64;

Query Match 25.2%; Score 888.5; DB 11; Length 474;
Best Local Similarity 35.0%; Pred. No. 2.4e-47;
Matches 234; Conservative 69; Mismatches 124; Indels 241; Gaps 18;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYFTGYMHVWVKSPGKGLGWIGRINPNNGVTLY 60
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
20 QVLOQSGELVKPGASVKISCKASGYAFSKSWNVVKKRRPGKGLGWIGRIFPCDGDTHY 79
QY 61 NOKFKDKATLVTDKSTTAYMELRSLTSEDSAVYYCARSTMITINYYMDYWGQGTSTVSS 120
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
80 SGKFOGKAKLTADKSSVTAFLQLTSLTSEDSAVYFCAVSDYGDY-FDDWGQCATVTSS 138
QY 121 AKTTPSVVPLAPGSAQAQNSMTVLCGLVKGYPEPVTYTNWNSGSLSSGVHHPFAVLQSD 190
DB :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :|||||
139 AKTTPSVVPLAPGCGDITGSSVTLGCLVKGYPEPVTYTNWNSGSLSSGVHHPFALLQSG 198
```



```
Query Match          24.1%; Score 850; DB 11; Length 234;
Best Local Similarity 76.9%; Pred. No. 2.3e-45;
Matches 163; Conservative 17; Mismatches 32; Indels 0; Gaps 0;

QY 460 IYVOTPTSLVSGDRVTITCKASQSVNDVAVYQKQPCQPKLLISYTSRYAGVPPDR 519
DB 22 IQLTQSPASLSASVGETVITCRASENIYSYLAWYQKQKSPQLLVYNAKTLADGVPSR 81

QY 520 FSGSGYGTDFTLTISSVQAEADRAVFCQDYNSPPTFGGKLEIKRADAAPTYSIFPPS 579
DB 82 FSGSRGTQFSLKINSLOPEDFGSYCQHSGIPFTFGSGTKLEIKRADAAPTYSIFPPS 141

QY 580 SLOLTSGGASVVCFLNFPKIDNVKWKIDGSRQNGVLNSWTDQDSKSTYSMSSTLTL 639
DB 142 SLOLTSGGASVVCFLNFPKIDNVKWKIDGSRQNGVLNSWTDQDSKSTYSMSSTLTL 201

QY 640 TKDEYERHNSYTCEATHKTSPIVKSFNRE 671
DB 202 TKDEYERHNSYTCEATHKTSPIVKSFNRE 233

RESULT 21
QY1205 PRELIMINARY; PRT; 473 AA.
AC Q91205;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TremBLrel. 21, Last annotation update)
DE Hypothetical 51.9 kDa protein.
GN AU044919.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010327; RAH10327.1;
DR MGI; MGI:2144967; AU044919.
DR InterPro; IPR000345; Cytochrome_c;
DR InterPro; IPR003006; Ig_MHC.
DR Pfam; PF00047; Ig; 3.
DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 473 AA; 51946 MW; CF625F008932AF12 CRC64;

Query Match          22.8%; Score 803; DB 11; Length 473;
Best Local Similarity 32.0%; Pred. No. 4.9e-42;
Matches 214; Conservative 74; Mismatches 138; Indels 242; Gaps 18;

QY 1 EVLOQSGDPLVKGASVKISKASGYSTGYMHVWVKSPKGLWIGRINPNNGVTLY 60
DB 20 EVLVESGGLVKGPGSRKLSAAGSTEDYGMHVRQAPKEGLEWVAYINSGSTTYY 79

QY 61 NQKFKDKATLTVDKSSITAYMELRSLSSESAVYICARSTMTITNYMDYWGQTSVTSS 120
DB 80 ADTVKGRFTISRDNAKNTFLQMTLSRSEDYAMCYARELWLR--IDYWGQGTITVSS 137

QY 121 AKTTPSPVPLAPGSAAGQNSMVTGLCLVKGYFPEPVTVTWNSGSLSSGVHTFPVAVLQSD 180
DB 138 AKTTPSPVPLAPGCGDTTGSSTVTLGLCLVKGYFPEPVTVTWNSGSLSSSVHTFPALLQSG 197

QY 181 LYTSSVTPSPSTWSPSETVTCNVAHPASTVKDKKIVPRDSGGPSEKSEINE-----KD 236
DB 198 LYTMSVTPSPSTWSPSTWSPQTSVAVHPASTVTDKKLEP---SGP---ISTINPCPKCE 251

QY 237 LRK--KSEQLGTALNKLKQIYYNSKAITSSSEKSADQFLNTLLKGFTHGFWYNDLLV 294
DB 252 CHKCAPNLEG---GPSVFIPFPNTKDV-----LMI 279
```

```
QY 295 DLGSTAATSEYEGSSVDLYGAYGYQCAGTGNKTKACMYGGVTLHDNRLTEKKVPINL 354
DB 280 SL-----TP-----283

QY 355 WIDGKQTTVPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVGRLIYFHSS 414
DB 284 ---KVTQVVD---VSEDDPDVQISWVFN--VEVHTA 313

QY 415 EGSTVSVDLFDAGQYPTDLLRIYRDNNTTISLSLSLYLTTSIVMTQTPTSLLSVAG 474
DB 314 QTQTHREDY-----NSTIR-----VYSA- 331

QY 475 DRVTITCKASQSVNDVAVYQKQPGSKLLISYTSRVAGVDPDRFSGSGYGTDFTLTIS 534
DB 332 -----LPIQHDMWSKEFKCKYN 350

QY 535 SYVAEDAAVYFCQDYNSPPTFGGKLEIKRADAAPTYSIFPPSSEQLTSGASVVCFL 594
DB 351 N-----KDLPSPIE---RTISKIKGLVRAPQVYILPPPAEQLSRKDVSUTCLV 395

QY 595 NNFPKIDINVKWKIDGSRQN-----GVLSWTDQDSKSTYSMSSTLTLTKDEYERHNS 649
DB 396 VGFNPGDISVENTSGHTEENYKDTAPVLDS-----DGSYFIYSKLDIKTSKNEKTD 448

QY 650 YTCSEATHK 657
DB 449 FSCNVRHE 456

RESULT 22
QY1217 PRELIMINARY; PRT; 268 AA.
AC Q85217;
DT 01-NOV-1998 (TremBLrel. 08, Created)
DT 01-NOV-1998 (TremBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Enterotoxin J.
GN SEJ.
OS Staphylococcus aureus.
OG Plasmid pIB485.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KS11410;
RX MEDLINE=99052098; PubMed=98350333;
RA Zhang S., Iandolo J.J., Stewart G.C.;
RT "The enterotoxin D plasmid of Staphylococcus aureus encodes a second
RT enterotoxin determinant (sej).";
RL FEMS Microbiol. Lett. 168:227-233(1998).
DR EMBL; AF053140; AAC78590.1;
DR HSP; P13163; IESF.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Plasmid.
SQ SEQUENCE 268 AA; 31230 MW; AC5F3546060ACE22 CRC64;

Query Match          21.8%; Score 768; DB 2; Length 268;
Best Local Similarity 58.3%; Pred. No. 3.4e-40;
Matches 141; Conservative 39; Mismatches 62; Indels 0; Gaps 0;

QY 228 KSEINEKDLRKSEKLGQGTALGNLKOIYYNSKAITSSSEKSADQFLNTLLKFGFTGHP 287
DB 27 KNETIKEKLNKHSSELSSITLNNLRIHYFFNEKGISEKIMTEQDFLDYTLLEKSFISIS 86

QY 288 WYNDLLVLGSLTAATSEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNRLTEE 347
DB 87 QYNDLLVQFDSKETNFKFGKQVDLYGSYGCSSGKPKNTKACMYGGVTLHENNQLYDT 146
```


Db 346 -----KCKVSNK-----GLP--- 355
QY 520 FSGSGYGTDTLTITSSVQAEADAAVFCQDYNPPTFGGKLEIKRADAAVTSIFPPS 579
Db 356 -----SSIEK-----TSKAGQPREPOVYTLPPS 380
QY 580 SROLTSGGASVVCFLNNFYPKIDINVKWKIDGSRQN-----GVLNSWTQDQSKDSTYSMS 634
Db 381 QEMTKNQVSLCLVKYGFPSDIADVESNGQPNKYKTPPVLDSE-----DGSFFLY 433
QY 635 STLTLTCKDEYERHNSYTCRATHK 657
Db 434 SRLTVDKSRWQGNVFCSCVMHE 456

RESULT 25
Q8TCD0
ID Q8TCD0 PRELIMINARY; PRT; 239 AA.
AC Q8TCD0;
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 26.2 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC022362; RAH22362.1;
KW Hypothetical protein.
SQ SEQUENCE 239 AA; 26234 MW; FAGEDC3A3B03871D CRC64;

Query Match 18.6%; Score 654.5; DB 4; Length 239;
Best Local Similarity 58.1%; Pred. No. 3.3e-33;
Matches 126; Conservative 34; Mismatches 52; Indels 5; Gaps 1;
QY 460 IVMTQPTSLVSGADRVITCKASQV-----SNDVAVQKQKPGSPKLLISYSSRYA 514
Db 22 VVMTQPLSLPVTLGQPASISCRSTQSLVSDGNTVLNWFQORPGQSRRLIVKYSNRDS 81
QY 515 GVPDRFSGSGYGTDTLTITSSVQAEADAAVFCQDYNPPTFGGKLEIKRADAAVTS 574
Db 82 GVPDRFSGSGYGTDTLTITSSVQAEADAAVFCQDYNPPTFGGKLEIKRADAAVTS 141
QY 575 IFPPSSEQLTSGGASVVCFLNNFYPKIDINVKWKIDGSRQNGVLSWTQDQSKDSTYSMS 634
Db 142 IFPPSDEQLKSGTASVVCFLNNFYPKIDINVKWKIDGSRQNGVLSWTQDQSKDSTYSLS 201
QY 635 STLTLTCKDEYERHNSYTCRATHKSTSPVKSFNRE 671
Db 202 STLTLSKADYERHKVYACEVTHQGLSSPVTKSFNRGE 238

RESULT 26
Q8VCX7
ID Q8VCX7 PRELIMINARY; PRT; 613 AA.
AC Q8VCX7;
DT 01-NAR-2002 (TReMBLrel. 20, Created)
DT 01-NAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Hypothetical 67.9 kDa protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SALIVARY GLAND;

RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC018315; AAH18315.1;
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig-cl.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; Igcl; 4.
DR SMART; SM00406; Ig; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67855 MW; 41A9384DD4C22862 CRC64;

Query Match 17.6%; Score 618.5; DB 11; Length 613;
Best Local Similarity 28.1%; Pred. No. 2e-30;
Matches 196; Conservative 90; Mismatches 211; Indels 201; Gaps 28;
QY 1 EVOLQSGPDLVKPGASVKISCKASGYFTGYMHVWKSPGKGLWIGRINPNNGVTLY 60
Db 20 QVOLQSGAELMKPGASVKISCKATGYTFSSWIEWKORPGHGLEIILPGSGSTNY 79
QY 61 NQKFKDKATLTVDKSTTAYMELSLTSDSAVYICARSTMITNMYDYGQSTVTSVS 120
Db 80 NEKFKGKATFTADTSSNTAYMQLSLTSDSAVYICAR--RLGRWYFDVWGAGTTVTSS 137
QY 121 AKTTPPSVYPLAP-GSAAOTNSMVLGCLVKGYFPEPVTVTV---NSGSLSSGVHTFPV 176
Db 138 ESQSPFNVPFLVSCESPLSKNLVAMGCLARDLPSTISFTWNYQNNTVIOGIRTFPTL 197
QY 177 LQSDLYTLSSVTV-PSS--TWPSETVTCNVAH-----PA---SSTKVDKKIYP 219
Db 198 RTGGKYLATSOVLLSPKLSILEGSDVLCVCKIHYGKNRDLHVPPIPAVAEMPNVNVFVP 257
QY 220 RDS-GGPSEKSEBEINEKDLRKSELOGTALGNLQIYYINYSKAITSEKSAQDLNTLL 278
Db 258 RDGFSGPAP-----RKSCL-----ICEATNFTPKPITVS----- 286
QY 279 FKGFTHGWYND-LLVDLGSTA--ATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGG 335
Db 287 -----WLKDGKLVESGFTDPVTIENKGSTPTQY----- 315
QY 336 VTLHNNRLTEKKVPINLWIDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLY 395
Db 316 -----KVIST---LTISEID-----WL 329
QY 396 NSDSFGKVGORGLIVPHSSEGSTVSVDLFDAGQGYPTLLRIYRDNTTISSTLSISLYL 455
Db 330 NLNVYTCRVDRHGLTFLKNVSTCA-----ASPSTDI----- 361
QY 456 YTTISIVMTQPTSLVSGADRVITCKASQVSN-----VANVQKQKPGSPKLLISYSSR 512
Db 362 ----LFTFTIPPADIFLSKANLTCVLSNLTATYETLNISWASQS-GEPLTKIKIMESH 416
QY 513 YAGVPDRFSGSGYGTDTLTITSSVQAEADAAVFCQDYNPPTFGGKLEIKR 566
Db 417 PNGT---FSAKG-----VASVCVEDNNRKEFVCTVTHRDLPSP-----QKFKISK 459
QY 567 ADAA-----PTVSIFPPSSSEQLT-SGGASVVCFLNNFYPKIDINVKWKIDGS--ERQNGVLN 619
Db 460 PNEVHKHPPAVYLLPPAREQLNRESATVTCVKGFSADISVQWLQRGOLLPOEKYVTS 519
QY 620 SWTDQDSKSTYSMSSTLTLTCKDEYERHNSYTCRATHK 657
Db 520 APMPEGAPGFYFTHSILTVEENNSGETYTCVWGHE 557

RESULT 27
Q96GA6
ID Q96GA6 PRELIMINARY; PRT; 614 AA.
AC Q96GA6;
DT 01-DEC-2001 (TReMBLrel. 19, Created)


```
Db 140 VTVSSGSASAPTLFPLVSCENSPSDTSS-VAVGCLAQDFLPDSITFSMKYKKNNSDISS-T 197
Qy 171 HTFPAVLQSDLYTLSSVTVPPSS---TWPSSETVTCNVAHAPASSTKVKKIIVPRDSCGPPSE 227
Db 198 RGFPVLRGGKYAATSOVLLPKQVMQGTDEHVCKVQHP-NGNKEKNVPLPVAIAELPK 256
Qy 228 KSEENEKDLRKSELOGTALGNLKAQIYYNSKAITSSSEKSAQDFLTNTLFLKGFETGHP 287
Db 257 VSVFVPPRD-----QATGFSP----- 285
Qy 288 WYNLDLLVGLGTAATSEYEGSSVDLYGAYGYQCAGTTPNKTACMYGGVTLHDNNRLTEE 347
Db 272 RKSKLIC-----QATGFSP----- 285
Qy 348 KVPINLWIDGKQ--TTVPIDKVKTSKKE-----VTVOELDLQARHYLHGKFL 394
Db 286 ROIQVSWLRGKQVGGVTTDQVQAEAKESGPTTYKVTSTLTIKESD-----W 333
Qy 395 YNSDSFGKQVQGLLIVFHSSEGSTVSYDLFDAQGYDPTLLRIYRDNTTISTSLISLY 454
Db 334 LSQSMFTCRVDHRLGTLFQONASSMCPD-----QDTAIRVF-----AIPPSFASIPL- 380
Qy 455 LYTTTSIVMTOTPTSLLSVAGDRVTITCKASQSVNDVAVYQKQSPKLLISYTSRYA 514
Db 381 --TKSTKLTLCLVTDL--TTYDSVVI-----SWTRQN-GEAVK---THNINISES 420
Qy 515 GVPDRFSSGSGVGTDTLTISSVQAEADAAYFCQODYNSPPTF-----GGTKLEIK 565
Db 421 HPNATFSAVG-----EASI--CEDDWSNGERFTCTVTHTDLPSPKQTIS 463
Qy 566 RADA---PTVSIFPPSSSEQLT-SGGASVVCFLNNFYPKDINVKKWIDGS--ERQNGVL 618
Db 464 RPKGVALHRPDVYLLPAREQLNLRRESATITCLVTGSPADVFQVMQRGQPLSPERYVT 523
Qy 619 NSWTDQDSKDYSTMSSTLTLTKDYEYERHNSYTCETHK 657
Db 524 SAPMPEQAPGRYFAHSLTVSEEWNTGETYTTCVVAHE 562

RESULT 33
Q9BQB8 PRELIMINARY; PRT; 597 AA.
AC Q9BQB8
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Unknown (protein for MGC:1905) (protein for MGC:1228).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC006180; AAH06180.1; -
DR EMBL; BC001872; AAH01872.1; -
DR HSSP; P01825; 7FAB.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003600; Ig_like.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; Ig; 2.
DR SMART; SM00407; Igcl; 4.
DR SMART; SM00406; Ig; 1.
```

```
DR SMART; SM00410; IG_like; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN 3.
SQ SEQUENCE 597 AA; 65300 MW; 2DAFAD50A6375851 CRC64;

Query Match 15.5%; Score 545; DB 4; Length 597;
Best Local Similarity 25.9%; Pred. No. 7.le-26;
Matches 182; Conservative 103; Mismatches 213; Indels 204; Gaps 29;

Qy 1 EVQLQQSGDPLVKGASVKISKASGYSTGYGMHMKVQSKGLGLEWIGRINPNNGVITY 60
Db 20 QVQLQQWAGALLKSETLSLTCGVYGGFSGYIWSWIRQPKGLGLEWIGIN-HSGITNY 78
Qy 61 NOKFRDKATLVDKSSSTTAYMELRSLTSEDSAVYYCARSTMIT-----NYVMDYWGQ 112
Db 79 NPSLKRSTVISVDTSKKQLSLKSLVNAADTAVYYCAR--VITRASPGTDRGMDWVGQ 136
Qy 113 GTSVTVSAAKTPPSVPL--APGSAATNSMTGLKGLVKGYFPFVTVW---NSGSL 167
Db 137 GTTVTVSSGSASAPTLFPLVSCENSPSDTSS-VAVGCLAQDFLPDSITFSMKYKKN 195
Qy 168 SGVHTFPAVLQSDLYTLSSSVTVPPSS---TWPSSETVTCNVAHAPASSTKVKKIIVPRD 224
Db 196 S-TRGFPVLRGGKYAATSOVLLPKQVMQGTDEHVCKVQHP-NGNKEKNVPLPVAIAEL 253
Qy 225 PSEKSEINEKDLRKSELOGTALGNLKAQIYYNSKAITSSSEKSAQDFLTNTLFLKGFET 284
Db 254 PPKVSVFVPPRD-----QATGFSP-----GFF- 268
Qy 285 GHPIYNDLLVGLGTAATSEYEGSSVDLYGAYGYQCAGTTPNKTACMYGGVTLHDNNRL 344
Db 269 GNPRKSLIC-----QATGFSP----- 285
Qy 345 TEKKKVPINLWIDGKQ--TTVPIDKVKTSKKE-----VTVOELDLQARHYLHGK 391
Db 286 ---ROIQVSWLRGKQVGGVTTDQVQAEAKESGPTTYKVTSTLTIKESD----- 332
Qy 392 FGLYNSDSFGKQVQGLLIVFHSSEGSTVSYDLFDAQGYDPTLLRIYRDNTTISTSLI 451
Db 333 --WLSQSMFTCRVDHRLGTLFQONASSMCPD-----QDTAIRVF-----AIPPSFASI 378
Qy 452 SLVLYTTSIVMTOTPTSLLSVAGDRVTITCKASQSVNDVAVYQKQSPKLLISYTS 511
Db 379 FL---TKSTKLTLCLVTDL--TTYDSVVI-----SWTRQN-GEAVK---THNINI 417
Qy 512 RYAGVDPDRFSSGSGVGTDTLTISSVQAEADAAYFCQODYNSPPTF-----GGTKL 562
Db 418 SEHPNATFSAVG-----EASI--CEDDWSNGERFTCTVTHTDLPSPKQ 460
Qy 563 EIKRADA---PTVSIFPPSSSEQLT-SGGASVVCFLNNFYPKDINVKKWIDGS--ERON 615
Db 461 TISRKGVALHRPDVYLLPAREQLNLRRESATITCLVTGSPADVFQVMQRGQPLSPK 520
Qy 616 GVLSNWDQDSKDYSTMSSTLTLTKDYEYERHNSYTCETHK 657
Db 521 YVTSAPMPEQAPGRYFAHSLTVSEEWNTGETYTTCVVAHE 562

RESULT 34
Q8WUX4 PRELIMINARY; PRT; 588 AA.
AC Q8WUX4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 64.4 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
```



```
Db 461 TISRPGVALHRPDVYLLPPAREQLNLRRESATITCLVTGSPADVFVQWMQRPGLSPEK 520
QY 616 GVLNSMTDQDSDKSTYSMSSTLTTLTKDEYERHNSYTCETHK 657
Db 521 YVTSAPMPEQAPGRYFAHSILIVSEEWNTGETYTCVVAHE 562
RESULT 36
Q96AA6 PRELIMINARY; PRT; 618 AA.
AC Q96AA6;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical 67.8 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LYMPH;
RA Strausberg R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC017356; AAH17356.1; -.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein; Immunoglobulin domain.
SQ SEQUENCE 618 AA; 67758 MW; 96DBD4C7C696E0A6 CRC64;

Query Match 15.3%; Score 540; DB 4; Length 618;
Best Local Similarity 25.9%; Pred. No. 1.5e-25;
Matches 182; Conservative 102; Mismatches 214; Indels 204; Gaps 29;

QY 1 EVLOQSGPDLVKPGASVKISKASGYSTFTGYMHVWKQSPGKGLIEWIGRINPNNGVTLY 60
Db 20 QVQLQWAGALLAPSETLSITCGYSGSYWIRPPGKGLIEWIGIN-HSGSTNY 78
QY 61 NQKFKDKATLVDSKSTTAYMELRLSTSDSAVYYCARSTMIT-----NYVMDYWGQ 112
Db 79 NPSLSRVTVISDTSKQLSLSSVNAADTAVYYCAR--VITRASPQTDGRYGMVWGQ 136
QY 113 GTSVTYSSAKTTPSYVPL--APGSAQTNSMTLGLCKLVGYPEPVTVTW---NSGSLS 167
Db 137 GTVTYSSGSASAPTLFPLVSCNSPDSSTSS--VAVGCLAQDFLPDSITFSWKYKNSDIS 195
QY 168 SGVHTFPVQLSDLYTLSSSVTPSS---TWPSETVTCNVNVAHPASSTKVDKKIVPRDSGG 224
Db 196 S-TRGFPSVLRGCKYAATSOVLLPSKDVMOGTDEHVVCVKQHP-NGNKEKNVPLVIAEL 253
QY 225 PSKSESEINEKDLRKSELOGTALGNLKOIYYNSKAITSSSEKSAQDFLNTLLFKGFFT 284
Db 254 PPKVSFVPPRD-----GFF- 268
QY 285 GHPWYNDLLVLDLSTAAATSEYEGSSVDLYGAYGYCAGGTPNKTACMGVTLHDNNRL 344
Db 269 GNPGRSKLLIC-----QATGFSP- 285
QY 345 TEKKVPIINLWIDGKQ--TTVPIDKVKTSKKE-----VTQELDLQARHYLHGK 391
Db 286 ---RQIQVSLRGEKQVSGVTTDQVAAEKESGPTTYKVTSLTIKESD----- 332
QY 392 FGLYNSDSFGKVGQRLIVFHSSEGSTVSYVDLDAQGVQPDTLRLRYRNTTISSTLSI 451
Db 333 --WLSQSMETCRVDHRGLTFQQAASSMCPVD-----QDTAIRVF---AIPPSFASI 378
QY 452 SLVLYTTSIVMTQPTPSLLVSAGDRVTITCKASQSVNDVAVYQKPGOSPKLLISYTS 511
Db 379 FL---TASTKLTCLVTDL--TTVDVSVTI-----SWTRQN-GEAVK---THTNI 417
```

```
QY 512 RYAGVDPDRFSGSGYGDFTLTISVQAEDAAYVFCQDYNPPTF-----GGGTKL 562
Db 418 SESHPNATFSAVG-----EASI--CEDDWSGERFCTCTVTHDLPSPKQ 460
QY 563 EIKRADAA-----PTVSIFPPSSSQLT--SGGASVCFLNPFYKDKINVKWKIDGS--ERQN 615
Db 461 TISRPGVALHRPDVYLLPPAREQLNLRRESATITCLVTGSPADVFVQWMQRPGLSPEK 520
QY 616 GVLNSMTDQDSDKSTYSMSSTLTTLTKDEYERHNSYTCETHK 657
Db 521 YVTSAPMPEQAPGRYFAHSILIVSEEWNTGETYTCVVAHE 562
RESULT 37
Q8WUK1 PRELIMINARY; PRT; 613 AA.
AC Q8WUK1;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 67.3 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020240; AAH20240.1; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003597; Ig_c1.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 5.
DR SMART; SM00409; IG; 2.
DR SMART; SM00407; IGC1; 4.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_3.
KW Hypothetical protein.
SQ SEQUENCE 613 AA; 67296 MW; 60C7F5950671E315 CRC64;

Query Match 15.3%; Score 537.5; DB 4; Length 613;
Best Local Similarity 25.4%; Pred. No. 2.1e-25;
Matches 176; Conservative 105; Mismatches 220; Indels 193; Gaps 26;

QY 1 EVLOQSGPDLVKPGASVKISKASGYSTFTGYMHVWKQSPGKGLIEWIGRINPNNGVTLY 60
Db 20 QVQLVESGGGVQPGKSLRLSCAASGFTSSYGMHWVRQAPGKGLWEWAVISYDGSNKYY 79
QY 61 NQKFKDKATLVDSKSTTAYMELRLSTSDSAVYYCARSTMITNYVMDYWGQTSVTSS 120
Db 80 ADSVKGRFTISRDNKNTLYLQMNLSRAEDTAVYYCAKDMSEGVETFDIMQGQTMVTSS 139
QY 121 AKTTPSPVPL--APGSAQTNSMTLGLCKLVGYPEPVTVTW---NSGSLSSGVHTFA 175
Db 140 GSASATLPLVLSVSCNSPDSSTSS--VAVGCLAQDFLPDSITFSWKYKNSDIS--TRGFPS 197
QY 176 VLQSDLYTLSSSVTPSS---TWPSETVTCNVNVAHPASSTKVDKKIVPRDSGGPSEKSEI 232
Db 198 VLRGGKYAATSOVLLPSKDVMOGTDEHVVCVKQHP-NGNKEKNVPLVIAELPPKVSFV 256
QY 233 NEKDLRKSELOGTALGNLKOIYYNSKAITSSSEKSAQDFLNTLLFKGFFTGHFWYNDL 292
Db 257 PPRD-----GFF-GNPRKSKL 271
QY 293 LVDLGSTAATSEYEGSSVDLYGAYGYCAGGTPNKTACMGVTLHDNNRLTEEEKVPI 352
Db 272 IC-----QATGFSP- 285
QY 353 NLWIDGKQ--TTVPIDKVKTSKKE-----VTQELDLQARHYLHGKFGLYNSDS 399
```

Db	286	SWLRKQVGVVTTQVQAEKESGPTTYKYVTSTLTIKESD-----WLSQSM	333
QY	400	FGKVGRLVIFVHSSRGSVSVYDLFDAQGYDPDTLLRIYRDNFTTISSTLSISLYLYTTS	459
Db	334	FTCRVDHRLGTLTQONASSMCVPD-----QDTARVF-----AIPPSASIFL---TKS	378
QY	460	IVNTQPTTLLVSAAGRVTITCKASOSVNDVAMVYQKPGQSPKLLISYTSRYAGVPDR	519
Db	379	TKLTLCLVTLTL--TTYDSVTI-----SWTRQN-GEAVK---THTNISESHPNAT	420
QY	520	FSGSGYGTDFTLTISSVQAEADAAYFCQDYNSPPTF-----GGCTKLEIKRADAA	570
Db	421	FSAVG-----EASI--CEDDWSNGERTCTVTHTDLPSPKQFISRPKGV	463
QY	571	-----PTVSIFPPSPSEQLT--SGGASVYVCFLLNFYPKDINVKWKIDGS--ERQNGVLNSWTD	623
Db	464	ALHRPDVYLLPPAREQLNRESATITCLVTGFSPADVFVQMQRGQPLSPKIVTSAPWP	523
QY	624	QDSKOSTYSMSSTLTLTQDEYERHNSYTCEATHK	657
Db	524	EPOAGRYFAHSILTYSEEWNTGETYTCVVAHE	557
RESULT 38			
Q924Q9			
ID	Q924Q9	PRELIMINARY;	PRT; 145 AA.
AC	Q924Q9;		
DT	01-DEC-2001 (TReMBLrel. 19, Created)		
DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TReMBLrel. 20, Last annotation update)		
DE	VH186.2-D-J-C mu protein (fragment).		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6;		
RA	Kozono Y., Kozono H., Azuma T.;		
RT	"Direct Estimation of Relative Affinity by Flow Cytometry Reveals		
RT	Affinity Maturation of B Cell Antigen Receptors in Response to (4-		
RT	Hydroxy-3-Nitrophenyl)Acetyl (NP).";		
RL	Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB067791; BAB63276.1; -		
DR	InterPro: IPR003006; Ig_MHC.		
DR	Pfam: PF00047; Ig; 1.		
FT	NON_TER 1		
FT	NON_TER 145 145		
SQ	SEQUENCE 145 AA; 16001 MW; 0F409EB09FA333D2 CRC64;		
Query Match			
Best Local Similarity 15.2%; Score 535; DB 11; Length 145;			
Matches 99; Conservative 15; Mismatches 17; Indels 0; Gaps 0;			
QY	1	EVOLQSGPDLVKPGASVKISKASGYSTGYVMHWKSPGKLEWIGRINPNNGVTLY	60
Db	1	QVLOQPGAEVLKPGASVKLSKASGYTETSYVMHWKQRPGRGLEWIGRIDPNSGCTKY	60
QY	61	NOKFKDKATLTVDKSTTAYMELRLSTSEDSAVYYCARSTMTITNYMDYWGQGTSTVTS	120
Db	61	NEKFKSKATLTVDKPSSTAYMQLSSLTSEDSAVYYCARSLITTYAYMDYWGQGTSTVTS	120
QY	121	AKTTPPSVYPL	131
Db	121	ESQSFNFVPL	131
RESULT 39			
Q91WT3			
ID	Q91WT3	PRELIMINARY;	PRT; 481 AA.
AC	Q91WT3;		
DT	01-DEC-2001 (TReMBLrel. 19, Created)		
DT	01-DEC-2001 (TReMBLrel. 19, Last sequence update)		

DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE	Hypothetical 52.0 kDa protein.
GN	AI893585.
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=COLON;
RA	Strausberg R.;
RL	Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; BC013488; AAH13488.1; -
DR	MGI; MGI:2144917; AI893585.
DR	InterPro; IPR003598; Ig_C2.
DR	InterPro; IPR003006; Ig_MHC.
DR	Pfam; PF00047; Ig; 4.
DR	SMART; SM00408; IGC2; 1.
DR	PROSITE; PS00290; IG_MHC; UNKNOWN_2.
KW	Hypothetical protein; Immunoglobulin domain.
SQ	SEQUENCE 481 AA; 52022 MW; 4EEB5C253038B718 CRC64;
Query Match	
Best Local Similarity 15.2%; Score 535; DB 11; Length 481;	
Matches 172; Conservative 79; Mismatches 168; Indels 246; Gaps 22;	
QY	1 EVQLQQSGPDLVKPGASVKISKASGYSTGYVMHWKQSPGKLEWIGRINPNNGVTLY 60
Db	20 QVLOQPGAEIVRPGASVKLSCKTSGYTFDYVMWVKQRPQGQLEWIGALDFDSTYSY 79
QY	61 NQKPKDKATLTVDKSTTAYMELRLSTSEDSAVYYCARSTMTITNYMDYWGQGTSTVTS 120
Db	80 NQKFKG--TTLTVDTSSTSSAYMLSSLTSEDSAVYFCARGPRDSSGY--YWGQGTTLTVSS 136
QY	121 AKTTPPSVYPLAPGSAQTNSMTLGLVKYGFPE--PVTVTW--NSGSLSSGVHTFPAVLQ 178
Db	137 EPAREPTIYPLTFPQALSSDPVI--IGCLIHDPSPGTMNVTMGKSGKDIITVFPFALAS 195
QY	179 SDLYTLSSSVTPSPSTWP--SETVTCNVAPASSTKVKIVPRDSDGSPSEKSEINEKDL 237
Db	196 GGRYTHSSQLTLPAVECPGESVKCSVQH-----DSNPVQELNV--- 234
QY	238 RKKSLEQGTALGNLKOIYYYNKSKAITSSSEKSAQOFITNTLLFKGFTTGHFWNDLLVDLG 297
Db	235 -----NCPGIC---SPPTTPPPPCQPSLS-----LQRPALDGL--LG 268
QY	298 STAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMGVGTLDHNNRLTEKKVPINLWID 357
Db	269 SDASIT-----CTLNG----- 279
QY	358 GKQTTVPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGKVGQGLIVFHSSEGS 417
Db	280 ----- 279
QY	418 TVSYDLFDAQGYDPTLLRIYRDNFTTISSTLSISLYLYTTSIVMTQPTSLVLSAGDRV 477
Db	280 -----LRDPEG-----AVFTWEP-----TGKDAV 299
QY	478 TITCKASQSVNDVAMVYQKPGQSPKLLISYTSRYAGVPDRFSGSGYGTDFTLTISVQ 537
Db	300 -----QKKAVQNSCGCY-----VSSVLPGCAERWNS---GASFCKCTVTHPE 338
QY	538 AEDAAVYFCQDYNSPPTFGGTTKLEIKRADAAAPTYSIRPPSSSEQLTSGG--ASVVCFLNN 596
Db	339 SDTLTGTTIAKVTNTFP-----PQVHLPPSPBELALNELVSLTCLVRA 382
QY	597 FYPKIDINVKWKIDGSRQNG-----VLNSWTDQDSKSTYSMSSTLTITKDEYERHNSYTC 652
Db	383 FNPKEVLVRW--LHGNEELSPESLYVFEPLKEPGEGATTLYTVSLVRSVAEIKWQGOQYSC 441
QY	653 EATHK 657
Db	442 MVGHE 446

Db	20	QVHLVQSGAEVMSGVSRVSCSTGSGYAFHTYSLIIVRWQAPGGLEWGMWISPSDNTFR	79
Qy	61	NQKFDRATLTVOKSSTAYMELRSLTSEDSAVVYCAR	112
Db	80	AKFQGRVTLTDTSTVVMELRSLRSDTAVYICARRYCYSVSSCONDYYYYYMDVMGK	139
Qy	113	GTSVTVSSAKTTPSPVYPLAPGSAAGTNSMWTGLCVKYGFP-EPVTVVMNSSLSSGVH	171
Db	140	GTTVTWSSASPTSPKVPPLSLCS-TPDGMNVWIACLVQGFQFPQELPSLVWSESGQGVJAR	198
Qy	172	TFPAV--LQSDLYTLSSSVTPSS-TWPSETVTCNAH---PASSTKVDKKIIVPRSGGP	225
Db	199	NFPSPQASGDLTYTTSQTLTPATCLAGKSVTCHVKHYNPQDVTVPCP-VPSTPPTP	257
Qy	226	SEKSEINEKDLRAKKSELOGTALGNLKOIYYNKAITSEKSAQOFLNTLLFKGFFTG	285
Db	258	SP-----STPTTSPSCCHPRLS-----LH	277
Qy	286	HPWYNLLVLGSLTAATSEYSGSVLDYGYQCGAGTPNKACWGVTLHDNRLT	345
Db	278	RPALEDLL--LGSEA-----NLCTLTG-----	298
Qy	346	EEKVPIINLWDGKQTTVPIDKYKTSKEVTVOELDQARHYLHGFEGLYNSDSFCGKVQ	405
Db	299	-----LRASG-----VTFWT	310
Qy	406	RGLIVFHSSEGSTVSYDLFDAQQGYPDTLLRIYRDNTTISSTLSISLVLYTTSIVMTOT	465
Db	299	-----LRASG-----VTFWT	310
Qy	466	PT----SLLVSAGDRVTTTCASOSVSNDA--WYQKPGQSPKLLISYTSRYAGVPDR	520
Db	311	PSSGKSAVQGPDPDRDLCCGYSVSVLSGCAEPW---NHGKTFTCTAAYPEK-----	359
Qy	521	SGSGYGTDFTLTTSVQAEADAAVYFCQODYNSPPTGGGTGLEIKRADAAPTVSIFFPSS	580
Db	360	-----TPLTATUS-----KSGNTF-----RPEVHLLPPPS	384
Qy	581	EQLTSGG--ASVVCFLNNFYPKDINVKWKIDGSRQNGVLNSWTDQSKD---STYSMSST	636
Db	385	EELALNELVTLCLARGFSFKDVLVRMLQGSQELPREKYLTVASRQEPSQGTTFVTSI	444
Qy	637	LTLTKDYEYRHNSYTCETHK	657
Db	445	LRVAEEDKKGDTFSCVMGHE	465

RESULT	45
Q924RI	
ID	PRELIMINARY; PRT; 145 AA.
AC	Q924RI:
DT	01-DEC-2001 (TtEMBLrel. 19, Created)
DT	01-DEC-2001 (TtEMBLrel. 19, Last sequence update)
DT	01-MAR-2002 (TtEMBLrel. 20, Last annotation update)
DE	VH186.2-D-J-C mu protein (Fragment).
OS	Mus musculus (Mouse).
QC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
RN	NCBI_TaxID=10090;
OX	[1]
RC	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6;
RA	Kozono Y., Kozono H., Azuma T.;
RT	"Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT	Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RL	Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL: AB067789; BAB63274.1;
DR	InterPro: IPR003006; Ig_MHC.
DR	Fram; PF00047; Ig; 1.
FT	NON_TER 1
FT	TER 145
FT	NON_TER 145

```

SQ  SEQUENCE  145 AA;  15979 MW;  0162D0A26C746C04  CRC64;

Query Match      14.7%;  Score 519;  DB 11;  Length 145;
Best Local Similarity 74.0%;  Pred. No. 4.3e-25;
Matches 97;  Conservative 15;  Mismatches 19;  Indels 0;  Gaps 0

Qy  1  EVQLQQSPDLVKPGASVKISKCKASGYSFTGYIMHWVKQSPGKGLEWIGRINPNNCVTILY 60
    :|:|:|:|  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db  1  QVQLQQPQAEELVPGASVKLSCKASGYTFTSYIMHWVKQRPGRGLEWIGRIDPNSGGTRY 60
    :|:|:|:|  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy  61  NOKFKKATLTVDKSSSTAYMELRSLTSEDSAVYVCARSTMTINVMYWGOGTSVTVSS 120
    :|:|:|:|  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db  61  NEKFKSKATLTVDKSSSTAYMQLSLTSEDSAVYVCARYDGSFSTAMDYWGOGTSVTVSS 120
    :|:|:|:|  :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|

Qy  121  AKTTPSPVYPL 131
    :|:|:|:|  :|:|:|:|
Db  121  ESQSPNPVFL 131
    :|:|:|:|  :|:|:|:|

```

RESULT	46
Q924Q3	
ID	PRELIMINARY; PRT; 146 AA.
AC	Q924Q3;
DT	01-DEC-2001 (TrEMBLrel. 19, Created)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE	VH186_2-D-J-C mu protein (Fragment).
OS	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=C57BL/6;
RA	Kozono Y., Kozono H., Azuma T.;
RT	"Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT	Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT	Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AB067797; BAB63282.1.;
DR	InterPro: IPR003006; Ig_MHC.
DR	Pfam; PF00047; ig; 1. 1
FT	NON_TER 1
FT	NON_TER 146 146
SO	SEQUENCE 146 AA; 16136 MW; CEA8DD6E1955807F CRC64;

```

Query Match          14.7%; Score 518.5; DB 11; Length 146;
Best Local Similarity 74.2%; Pred. NO. 4.7e-25;
Matches 98; Conservative 15; Mismatches 18; Indels 1; Gaps 1

QY 1 EVQLQQSPDLVKPCASVKISCKASYGFTGYMHHWKQSPCKGLEWIGRIPNNGVTLY 60
   :||||| : :||||| : :||||| : :||||| : :||||| : :||||| : :||
Db 1 QVQLQQGAEGLVKPCASVKLSCKASGYFTFTYMHHWKQRPGRGLEWIGRIDPNSGGTKY 60
   :||||| : :||||| : :||||| : :||||| : :||||| : :||||| : :||

QY 61 NQKFKDKATLTVDKSGTAYMELRLTSDSDSAVYYCARSTM-ITNVMDYWGQGTSVTVS 119
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 61 NEKFKSKATLTVDKFSSTAYMQLSLTSDSDSAVYYCARSLYDYGDAYMDYWGQGTSVTVS 120
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 120 SAKTTPPSVYPL 131
   I : :|||
Db 121 SESQSFNPVFL 132

RESULT 47
Q924P7 PRELIMINARY; PRT; 145 AA.
ID Q924P7;
AC Q924P7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE VH186.2-D-J-C mu protein (fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```



```
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB065918; BAB63934.1; -.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig: 1.
FT NON_TER 1
FT NON_TER 145
SQ SEQUENCE 145 AA; 15988 MW; FB73958704796C9A CRC64;

Query Match 14.6%; Score 515; DB 11; Length 143;
Best Local Similarity 72.5%; Pred. No. 7.7e-25;
Matches 95; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYSTGYMHVWKSPGKGLWIGRINPNGVTLY 60
Db 1 QVLOQPGALVPGSVKLSCKASGYTFTSYMHVWKORPGKGLWIGRIDPNSGGTKY 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NQKFKDKATLTVDKSSTTAYMELRLSITSDSAVYYCARSTMTITNYMDYWGQTSVTSS 120
Db 61 NEKFKSKATLTVDKPSSTAYMOLSLTSDSAVYYCARSGSFYAMDYWGQTSVTSS 120
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 AKTTPPSVYPL 131
Db 121 ESQSFNPVPL 131
:|||||:

RESULT 48
Q924R0 PRELIMINARY: PRT; 143 AA.
AC Q924R0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-WAR-2002 (TrEMBLrel. 20, Last annotation update)
DE VH186.2-D-J-C mu protein (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Kozono Y., Kozono H., Azuma T.;
RT "Direct Estimation of Relative Affinity by Flow Cytometry Reveals
RT Affinity Maturation of B Cell Antigen Receptors in Response to (4-
RT Hydroxy-3-Nitrophenyl)Acetyl (NP).";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB067790; BAB63275.1; -.
DR InterPro: IPR003598; Ig_C2.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig: 1.
DR SMART: SM00408; IGC2; 1.
KW Immunoglobulin domain.
FT NON_TER 1
FT NON_TER 143
SQ SEQUENCE 143 AA; 15868 MW; 139B2E9668B1E07F CRC64;

Query Match 14.6%; Score 514; DB 11; Length 143;
Best Local Similarity 74.0%; Pred. No. 8.7e-25;
Matches 97; Conservative 15; Mismatches 17; Indels 2; Gaps 1;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYSTGYMHVWKSPGKGLWIGRINPNGVTLY 60
Db 1 QVLOQPGALVPGASVKLSCKASGYTFTSYMHVWKORPGKGLWIGRIDPNSGGTKY 60
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NQKFKDKATLTVDKSSTTAYMELRLSITSDSAVYYCARSTMTITNYMDYWGQTSVTSS 120
```

```
Db 61 NEKFKSKATLTVDKPSSTAYMOLSLTSDSAVYYCAR--WDEDYAMDYWGQTSVTSS 118
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 121 AKTTPPSVYPL 131
Db 119 ESQSFNPVPL 129
:|||||:

RESULT 49
Q8WU38 PRELIMINARY: PRT; 573 AA.
AC Q8WU38;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Hypothetical 63.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC021276; AAH21276.1; -.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003006; Ig_MHC.
DR Pfam: PF00047; Ig: 4.
DR SMART: SM00409; IG: 1.
DR SMART: SM00407; IGC1; 3.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_2.
KW Hypothetical protein.
SQ SEQUENCE 573 AA; 62967 MW; FD072344033AC530 CRC64;

Query Match 14.6%; Score 514; DB 4; Length 573;
Best Local Similarity 27.8%; Pred. No. 5.6e-24;
Matches 157; Conservative 85; Mismatches 201; Indels 122; Gaps 18;

QY 1 EVLOQSGDPLVKPGASVKISCKASGYSTGYMHVWKSPGKGLWIGRINPNGVTLY 60
Db 20 EVQLVESGGGLVQPGSRSLRSLCAASGFTFDYAMHWVRQAPGKLEWVSGISNNGSIGY 79
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 61 NQKFKDKATLTVDKSSTTAYMELRLSITSDSAVYYCAR----STMTITNYMDYWGQTSV 116
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 80 ADSVKGRFTISRDNAKNSLYQMNSLRRAEDTALYYCAKHGSGSYIGYYGMVWGQTTV 139
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 117 TVSSAKTTPPSVYPLAPGSA-AQTNSMVTLGCLVKGYFPEPTVTWNSGSLSCVHTFPA 175
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 140 TVSSATKAPDVEFIISGCRHPKDNSPVVLACLTGYHPTSVTVTWYWGQSQPQRTFPE 199
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 176 VLQSDLYTLSSS-VTVPSSTWSPSETVTCNVAPHAS-----STKVDKKIVP---- 219
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 200 IQRRDSYMTSSOLSTPLQWRQGEYKCVVQHTASKSKKEIFRPESPKAQASVPPTAQP 259
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 220 -----RDSGGPSEKSEEINEKDLRKKSSELOQTALGNLKI--YYNSKA 261
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 260 QAEGSLAKATTAPATTATNTGRGGEKKKEKEBEQERETKTPECPSTHTQPLGVYLLTPA 319
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 262 ITSSKSAQDFLTNTLLFKGFFTHPWYNDLLVDLGLSTAATSEYEGSSVDLYGAYGYQC 321
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 320 V-----QDLWLRLDKATFTCFVYVGS-----DLKDAHLTWEV 349
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 322 AGGTGNKTKACMGV-----TLHDNNRLTEKK--VPINLWIDGKQTVPIDRVK-TSKK 373
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 350 AGKVPT-----GGVEGLLERHNSGQSQHSRLTLPRLSNAGTSVTCTLNHPSLPQR 403
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 374 EVTVQELDLQARHYLHGKFGLYNSDFGGKVGQRLGVFIHSS--GSTVSYDLFDAQGGYPD 432
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 404 LMALEPAAQA-----PVKLSLNLASSDPPPEAASWLLCEVSGFSPP 445
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:52:01 ; Search time 16.9923 Seconds
(without alignments)
1827.149 Million cell updates/sec

Title: US-09-900-766-2

Perfect score: 1218

Sequence: 1 SEKSEINEKDKRKSELOG.....RDNTTISSTLSISLYLTT 233

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDS2/qcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1113	91.4	245	18	Staphylococcus ent
2	1107	90.9	257	22	Peptide sequence f
3	1093	89.7	230	22	Staphylococcus aur
4	1091	89.6	245	18	Staphylococcus ent
5	1048	86.0	230	14	Staphylococcus ent
6	1044	85.7	230	12	Staphylococcus aur
7	1035	85.0	230	23	Staphylococcus ent
8	960	78.8	233	12	Staphylococcus ent
9	948	77.8	257	22	Peptide sequence f
10	944	77.5	233	18	Staphylococcus ent

11	942	77.3	233	14	AA45011	Staphylococcal ent
12	941	77.3	233	18	AAW35373	Staphylococcus ent
13	941	77.3	233	22	AAW67338	Staphylococcus aur
14	935	76.6	233	23	ABB76234	Staphylococcus aur
15	925	75.9	257	23	ABB79501	Staphylococcal ent
16	921	75.6	233	21	AA54463	Amino acid sequenc
17	921	75.6	233	23	ABB79502	Staphylococcal ent
18	908	74.5	257	21	AAW70102	Staphylococcal ent
19	903	74.1	233	21	AAW70103	Mutant Staphylococ
20	605	49.7	228	14	AA45013	Staphylococcal ent
21	605	49.7	228	22	AAW67340	Staphylococcus aur
22	605	49.7	228	23	ABB76236	Staphylococcus aur
23	597	49.0	228	12	AAW13205	Staphylococcal ent
24	370	30.4	91	18	AAW24299	Staphylococcus aur
25	291.5	23.9	259	23	ABP29357	Streptococcus poly
26	287.5	23.6	239	20	AAV06254	Staphylococcal gro
27	287.5	23.6	239	20	AAV06253	Staphylococcal gro
28	282.5	23.2	239	20	AAV06256	Staphylococcal gro
29	280.5	23.0	239	20	AAV06255	Staphylococcal gro
30	279.5	22.9	239	20	AAV06252	Staphylococcal gro
31	277.5	22.8	238	14	AA45016	Staphylococcal ent
32	275.5	22.6	238	22	AAW67343	Staphylococcus aur
33	275.5	22.6	238	23	ABB76239	Staphylococcus aur
34	273.5	22.5	228	22	AAW63856	Amino acid sequenc
35	272.5	22.4	238	12	AAW13208	Staphylococcal ent
36	271.5	22.3	251	18	AAW12153	Streptococcus pyog
37	270.5	22.2	221	12	AAW13209	Streptococcus pyog
38	270.5	22.2	221	14	AAW45017	Staphylococcal ent
39	270.5	22.2	221	23	ABB76240	Staphylococcus pyo
40	270.5	22.2	251	18	AAW12150	Streptococcus pyog
41	270.5	22.2	251	18	AAW12151	Streptococcus pyog
42	268.5	22.0	251	18	AAW12154	Streptococcus pyog
43	268.5	22.0	251	18	AAW12148	Streptococcus pyog
44	268.5	22.0	251	18	AAW12147	Streptococcus pyog
45	268.5	22.0	251	18	AAW12146	Streptococcus pyog

ALIGNMENTS

RESULT 1
AAW35375
ID AA35375 standard; peptide: 245 AA.
XX
AC AA35375;
XX
DT 20-APR-1998 (first entry)
XX
DE Staphylococcus enterotoxin SEE modified superantigen.
XX
KW SEE; SEA; Staphylococcus enterotoxin; superantigen; conjugate;
XX treatment; cancer; infection; autoimmune disease; antibody; modified.
XX
OS Staphylococcus sp.
XX
FH Key Location/Qualifiers
XX Misc-difference 20
FT /label= R20G
FT /note= "wild-type Arg is replaced by Gly"
FT Misc-difference 21
FT /label= N21T
FT /note= "wild-type Asn is replaced by Thr"
FT Misc-difference 24
FT /label= S24G
FT /note= "wild-type Ser is replaced by Gly"
FT Misc-difference 27
FT /label= R27K
FT /note= "wild-type Arg is replaced by Lys"
XX W09736932-A1.
XX 09-OCT-1997.
XX

AAW35374 standard; peptide; 245 AA.

AAW35374;

20-APR-1998 (first entry)

Staphylococcus enterotoxin SEE wild-type superantigen.

SEE; SEA; Staphylococcus enterotoxin; superantigen; conjugate; treatment; cancer; infection; autoimmune disease; antibody.

Staphylococcus sp.

Key Location/Qualifiers

Misc-difference 20 /note= "can be mutated at this position"

Misc-difference 21 /note= "can be mutated at this position"

Misc-difference 24 /note= "can be mutated at this position"

Misc-difference 27 /note= "can be mutated at this position"

WO9736932-A1.

09-OCT-1997.

26-MAR-1997; 97WO-SE00537.

12-AUG-1996; 96US-0695692.

29-MAR-1996; 96SE-0001245.

(PHAA) PHARMACIA & UPJOHN AB.

Abrahmsen L, Antonsson P, Bjoerk P, Dohlsen M; Forsberg G, Hansson J, Kalland T;

WPI; 1997-503052/46.

Conjugate of target seeking moiety and modified superantigen - useful for activating the immune system to treat cancer, viral infections, parasitic infestations and autoimmune diseases

Claim 4; Pages 38-39; 58pp; English.

This is the wild-type Staphylococcus enterotoxin SEE superantigen. This SEE superantigen can be modified to be used in a novel conjugate. The novel conjugate comprises a target seeking moiety and a modified wild type superantigen. The modified superantigen retains its ability to activate a subset of T cells, even though 1 or more wild-type amino acid residues in at least 1 region which functions in determining binding to T cell receptor (TCR) and activation of a subset of T cells has/have been replaced. Such a modified superantigen can optionally be used as part of a conjugate with a target seeking moiety, for activating the immune system to treat a mammalian disease. A pharmaceutical composition can be prepared comprising a modified antibody (preferably a Fab fragment fused to a peptide moiety providing activation of T cells in Vbeta specific manner) in which cysteines providing for interchain cysteine linkages in the native antibody have been replaced (preferably by serine residues) to prohibit cysteine formation. The modified wild-type superantigen is used for treating cancer, viral infections, parasitic infestations and autoimmune disease. The modified wild type superantigen has a lower immunogenicity and reactivity with neutralising antibodies and has fewer side-effects when used as a drug, compared to wild type superantigen.

Sequence 245 AA;

Query Match 89.6%; Score 1091; DB 18; Length 245;

Best Local Similarity 85.3%; Pred. No. 1.2e-102;

Matches 209; Conservative 9; Mismatches 15; Indels 12; Gaps 1;

QY 1 SEKSEEINEKDLRKSELOGTALGNLKOIYYNKAITSSEKSADQFTNTLLKFGFTG 60
|||||

AAW35374 standard; peptide; 230 AA.

AAW35374;

23-APR-2001 (first entry)

Staphylococcus aureus enterotoxin E protein.

Tumour; cancer; immune; enterotoxin.

Staphylococcus aureus.

US6180097-B1.

30-JAN-2001.

30-OCT-1998; 98US-0183437.

31-JAN-1994; 94US-0189424.

19-JUN-1995; 95US-0491746.

03-OCT-1989; 89US-0416530.

17-JAN-1990; 90US-0466577.

17-JAN-1991; 91WO-US00342.

01-JUN-1992; 92US-0891718.

02-MAR-1993; 93US-0025144.

(TERM/) Terman D S.

Terman DS;

WPI; 2001-158657/16.

Tumor cell capable of stimulating antitumor immune reactivity in vitro or in vivo comprises exogenous nucleic acids encoding a superantigen and a costimulatory molecule -

Disclosure; Fig 2; 16pp; English.

The present invention relates to a tumour cell capable of stimulating antitumor immune reactivity in vitro or in vivo contains and expresses an exogenous nucleic acid molecule encoding a superantigen or its active fragment and an exogenous nucleic acid molecule encoding a costimulatory molecule that activates T cells in conjunction with an antigenic stimulus. The invention may be used for cancer therapy by stimulating an anticancer immune response in vivo or ex vivo.

Sequence 230 AA;

Query Match 89.7%; Score 1093; DB 22; Length 230;

Best Local Similarity 89.6%; Pred. No. 6.7e-103;

Matches 206; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 4 SEEINEKDLRKSELOGTALGNLKOIYYNKAITSSEKSADQFTNTLLKFGFTGHPW 63
|||||

Db 1 SEEINEKDLRKSELOGTALGNLKOIYYNKAITSSEKSADQFTNTLLKFGFTGHPW 60
|||||

QY 64 YNDLLVDLGSKATNRYKGGKVDLGGYGYOCAGTGNKTKACMYGGVTLHDNNRLTEK 123
|||||

Db 61 YNDLLVDLGSKATNRYKGGKVDLGGYGYOCAGTGNKTKACMYGGVTLHDNNRLTEK 120
|||||

QY 124 KVPINLWIDGKQTPIDKVKTSKREVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRL 183
|||||

Db 121 KVPINLWIDGKQTPIDKVKTSKREVTVQELDLQARHYLHGKFGLYNSDSFGGKVQRL 180
|||||

QY 184 IVFHSSEGSTVSYDLFDAGQGVPTDLLRIYRDNNTTSISLSILYLYTT 233
|||||

Db 181 IVFHSSEGSTVSYDLFDAGQGVPTDLLRIYRDNNTTSISLSILYLYTT 230
|||||

RESULT 4

AAW35374

Db 1 SEKSEINEKDLRKKSELQARNALSLRQIYYNEKAITENKESDDQFLNTLLFKGFFTG 60
Qy 61 HPWYNLLVDLGLGTAATSEYEGSSVDLYCAYYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 61 HPWYNLLVDLGLGSKDATNKYKGGKVDLYGAYYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
Qy 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHLYHGKFGLYNSDSFGGKVQ 180
Db 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHLYHGKFGLYNSDSFGGKVQ 180
Qy 181 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTINSENHIDL 228
Db 181 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTINSENHIDL 240
Qy 229 YLYTT 233
Db 241 YLYTT 245
RESULT 5
ID AAR45012 standard; protein; 230 AA.
XX AAR45012;
AC AAR45012;
XX 08-JUN-1994 (first entry)
XX Staphylococcal enterotoxin SEE.
DE Staphylococcal enterotoxin SEE.
XX Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
KW auto-immune disease; toxicity; Protein A; perfusion system.
XX Staphylococcus aureus.
OS Staphylococcus aureus.
FH Key Location/Qualifiers
FT Misc-difference 120 /note= "Given in the specification as J, no further
FT details given"
FT Misc-difference 121 /note= "Given in the specification as J, no further
FT details given"
FT Misc-difference 123 /note= "Given in the specification as O, no further
FT details given"
FT Misc-difference 124 /note= "Given in the specification as U, no further
FT details given"
XX WO9324136-A.
XX 09-DEC-1993.
XX 01-JUN-1993; 93WO-US05213.
XX 01-JUN-1992; 92US-0891718.
XX (STON/) STONE J L.
XX (TERM/) TERMAN D S.
XX Stone JL, Terman DS;
PI WPI: 1993-405418/50.
XX WPI: 1993-405418/50.
XX Use of staphylococcal enterotoxin(s) and homologues - for
PT treating cancer in a patient or for the treatment of auto-immune
PT diseases
XX Disclosure: Fig 1; 90pp; English.
XX The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
CC which may be used in the methods of the invention for treating cancer
CC in a patient. These SEs, and homologues of them, can be used as
CC tumouricidal agents for treating cancers and autoimmune disease.

CC They exhibit tumouricidal activity and toxicity identical to that
CC observed for the Protein A perfusion system. They may be administered
CC by i.v. injection.
XX
SQ Sequence 230 AA;
Query Match 86.0%; Score 1048; DB 14; Length 230;
Best Local Similarity 85.7%; Pred. No. 2.6e-98;
Matches 197; Conservative 12; Mismatches 21; Indels 0; Gaps 0;
Qy 4 SEETNEKDLRKKSELQGTALGNLKOIYYNSKAITSEKSAQDQFLNTLLFKGFFTGHPW 63
Db 1 SEETNEKDLRKKSELQARNALSLRQIYYNEKAITENKESDDQFLNTLLFKGFFTGHPW 60
Qy 64 YNDLLVDLGLGTAATSEYEGSSVDLYCAYYGYOCAGGTPNKTACMYGGVTLHDNNRLTEK 123
Db 61 YNDLLVDLGLGSKDATNKYKGGKVDLYGAYYGYOCAGGTPNKTACMYGGVTLHDNNRLTEK 120
Qy 124 KVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHLYHGKFGLYNSDSFGGKVQ 183
Db 121 VXXBKWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHLYHGKFGLYNSDSFGGKVQ 180
Qy 184 IVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTINSENHIDL 233
Db 181 IVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTINSENHIDL 230
RESULT 6
ABB76235
ID ABB76235 standard; Protein; 230 AA.
XX ABB76235;
AC ABB76235;
XX 09-AUG-2002 (first entry)
XX Staphylococcus aureus enterotoxin E.
DE Enterotoxin E; SEE; superantigen; antigen; tumour; cancer;
KW antitumour; therapy.
XX Staphylococcus aureus.
OS Staphylococcus aureus.
FH Key Location/Qualifiers
FT Misc-difference 120 /note= "given as 'J' in the specification"
FT Misc-difference 121 /note= "given as 'J' in the specification"
FT Misc-difference 123 /note= "given as 'O' in the specification"
FT Misc-difference 124 /note= "given as 'U' in the specification"
FT Misc-difference 125 /note= "given as 'V' in the specification"
XX US2002051765-A1.
XX 02-MAY-2002.
XX 19-DEC-2000; 2000US-0741503.
XX 31-JAN-1994; 94US-0189424.
XX 19-JUN-1995; 95US-0491746.
XX 03-OCT-1989; 89US-0416530.
XX 17-JAN-1990; 90US-0466577.
XX 17-JAN-1991; 91WO-US00342.
XX 01-JUN-1992; 92US-0891718.
XX 02-MAR-1993; 93US-0025144.
XX (TERM/) TERMAN D S.
XX Terman DS;
XX WPI: 2002-415198/44.
DR

XX Reagent for treating cancer without the need for e.g. radiotherapy,
PT comprises a specific V beta subset of T cells sensitized to a growing
PT tumor and stimulated with superantigens -
XX
PS Disclosure; Fig 2; 17pp; English.
XX The present sequence is the protein sequence of enterotoxin E
CC (SEE) of Staphylococcus aureus. Similarity is shown, in several
CC stretches of sequence, between staphylococcal enterotoxins,
CC streptococcal pyrogenic exotoxins and staphylococcal exfoliative
CC toxins (see ABB76234-44). In the present invention, synthetic
CC polypeptides useful in tumour therapy and in blocking or destroying
CC autoreactive T and B lymphocyte populations are characterised by
CC substantial structural homology to staphylococcal enterotoxin A and
CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with
CC statistically significant sequence homology and similarity (Z value
CC of Lipman and Pearson algorithm in Monte Carlo analysis exceeding
CC 6) to include alignment of cysteine residues and similar hydrophathy
CC profiles. These superantigens are used to treat solid tumours,
CC including their metastases, without radiation, surgery or
CC standard chemotherapeutic agents. A claimed method of human cancer
CC treatment involves contacting haematopoietic cells from a patient
CC with one or more superantigens ex vivo to generate stimulated cells,
CC selecting a specific V beta subset of cells, and reintroducing
CC these cells into the patient to induce an in vivo therapeutic,
CC tumouricidal reaction.
XX
SQ Sequence 230 AA;
Query Match 85.7%; Score 1044; DB 23; Length 230;
Best Local Similarity 85.7%; Pred. No. 6.5e-98;
Matches 197; Conservative 11; Mismatches 22; Indels 0; Gaps 0;
Qy 4 SEINEKDLRKSELGQTALGNLKOIYYNSKAITSSSEKSDQFLNTLLPKGFTGHPW 63
Db 1 SEINEKDLRKSELRNALSNLRQIYYNEKAKTENKESHDDQFLQHTILFKGFTGHPW 60
Qy 64 YNDLLVDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKACMGYGVTLHDNNRLTEEK 123
Db 61 YNDLLVDGSKDATNKYKGGKVDLYGAYGYOCAGGTPNKACMGYGVTLHDNNRLTEEX 120
Qy 124 KVPINLWIDGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFGLYNSDSFGGKVQRGL 183
Db 121 XVXXKWDIGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFGLYNSDSFGGKVQRGL 180
Qy 184 IVFHSSEGSTVSYDLFDAQQGYPTDPLLRIYRDNNTTISLSISLYLYTT 233
Db 181 IVFHSSEGSTVSYDLFDAQQGYPTDPLLRIYRDNNTTINSENMHIDIYLYTT 230
Qy 184 IVFHSSEGSTVSYDLFDAQQGYPTDPLLRIYRDNNTTISLSISLYLYTT 233
Db 181 IVFHSSEGSTVSYDLFDAQQGYPTDPLLRIYRDNNTTINSENMHIDIYLYTT 230
RESULT 7
AAR13204
ID AAR13204 standard; Protein; 230 AA.
XX AAR13204;
AC AAR13204;
XX 15-OCT-1991 (first entry)
DT Staphylococcal enterotoxin E.
DE Staphylococcus aureus.
XX SEE; cancer treatment; pyrogen; tumouricide.
KW Staphylococcus aureus.
OS Staphylococcus aureus.
PN WO9110680-A.
XX 25-JUL-1991.
PD 17-JAN-1991; 91WO-US00342.
XX 17-JAN-1991; 91WO-US00342.
PF 17-JAN-1991; 90US-0466577.
PR 17-JAN-1990; 90US-0466577.
XX

PA (TERM/) TERMAN D S.
XX Terman DS;
PI WPI; 1991-237984/32.
XX
DR WPI; 1991-237984/32.
XX
PT Treating cancer with enterotoxin from Staphylococcus aureus -
PT administered by IV injection, having same tumoricidal activity
PT as Staphylococcal protein A without potential toxic reactions
XX
PS Disclosure; Fig 1; 74pp; English.
XX SEE was isolated and purified from S.aureus. It can be used for
CC treating cancer, activating cytokine mediators and procoagulant
CC systems, augmenting natural killer cell cytotoxicity, etc. The
CC enterotoxin can be administered intravenously, optionally with
CC ibuprofen to attenuate toxic reaction to SEE. Synthetic
CC polypeptides having structural homology to Staphylococcal exotoxins
CC are claimed, provided the homology includes statistically
CC significant sequence homology, alignment of Cysteine residues and
CC similar hydrophathy profiles.
CC See AAR13203-R13211.
XX
SQ Sequence 230 AA;
Query Match 85.0%; Score 1035; DB 12; Length 230;
Best Local Similarity 84.3%; Pred. No. 5.4e-97;
Matches 194; Conservative 14; Mismatches 22; Indels 0; Gaps 0;
Qy 4 SEINEKDLRKSELGQTALGNLKOIYYNSKAITSSSEKSDQFLNTLLPKGFTGHPW 63
Db 1 SEINEKDLRKSELRNALSNLRQIYYNEKAKTENKESHDDQFLQHTILFKGFTGHPW 60
Qy 64 YNDLLVDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKACMGYGVTLHDNNRLTEEK 123
Db 61 YNDLLVDGSKDATNKYKGGKVDLYGAYGYOCAGGTPNKACMGYGVTLHDNNRLTEEX 120
Qy 124 KVPINLWIDGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFGLYNSDSFGGKVQRGL 183
Db 121 XVQXBKWDIGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFGLYNSDSFGGKVQRGL 180
Qy 184 IVFHSSEGSTVSYDLFDAQQGYPTDPLLRIYRDNNTTISLSISLYLYTT 233
Db 181 IVFHSSEGSTVSYDLFDAQQGYPTDPLLRIYRDNNTTINSENMHIDIYLYTT 230
RESULT 8
AAR13203
ID AAR13203 standard; protein; 233 AA.
XX AAR13203;
AC AAR13203;
XX 15-OCT-1991 (first entry)
DT Staphylococcal enterotoxin A.
DE Staphylococcus aureus.
XX SEA; cancer treatment; pyrogen; tumouricide.
KW Staphylococcus aureus.
OS Staphylococcus aureus.
PN WO9110680-A.
XX 25-JUL-1991.
PD 17-JAN-1991; 91WO-US00342.
XX 17-JAN-1991; 90US-0466577.
PR 17-JAN-1990; 90US-0466577.
XX (TERM/) TERMAN D S.
PA Terman DS;
PI WPI; 1991-237984/32.
XX
DR WPI; 1991-237984/32.

```
XX Treating cancer with enterotoxin from Staphylococcus aureus -
PT administered by IV injection, having some tumoricidal activity
PT as Staphylococcal protein A without potential toxic reactions
XX
PS Disclosure; Fig 1; 74pp; English.
XX
XX SEA was isolated and purified from S.aureus. It can be used for
CC treating cancer, activating cytokine mediators and procoagulant
CC systems, augmenting natural killer cell cytotoxicity, etc. The
CC enterotoxin can be administered intravenously, optionally with
CC ibuprofen to attenuate toxic reaction to SEA. Synthetic
CC polypeptides having structural homology to Staphylococcal
CC are claimed, provided the homology includes statistically
CC significant sequence homology, alignment of Cysteine residues and
CC similar hydropathy profiles.
CC See also AAR13204-R13211.
XX
SQ Sequence 233 AA;
Query Match 78.8%; Score 960; DB 12; Length 233;
Best Local Similarity 77.3%; Pred. No. 2.4e-89;
Matches 180; Conservative 21; Mismatches 32; Indels 0; Gaps 0;
Qy 1 SEKSEINEKDLRKSELGQTALGNLKOIYYNYSKAITSSSEKSAQDFTLTLFKGFFTG 60
Db 1 SEKSEINEKDLRKSELGQTALGNLKOIYYNYSKAITSSSEKSAQDFTLTLFKGFFTG 60
Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 61 HSWYNDLLVDLKGSKDIDVYKGGKVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
Qy 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
Db 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
Qy 181 RGLIVFHSSEGSTVSYDLFDAAGQYSPDFTLLRIYRDNNTTISSTLSLSISLYLTT 233
Db 181 RGLIVFHTSTEPSVNYDLFDAAGQYSNTLLRIYRDNKNTINSENMHIDIYLYTS 233
RESULT 9
AAU14104
ID AAU14104 standard; peptide: 257 AA.
XX
XX AAU14104;
XX
XX 21-NOV-2001 (first entry)
DT
DE Peptide sequence from Staphylococcus aureus enterotoxin A.
XX
XX Anti-retroviral; DP178-like; DP107-like; enterotoxin A;
KW antifusogenic; antiviral; HIV transmission.
XX
OS Staphylococcus aureus.
XX
XX WO200151673-A2.
PN
XX 19-JUL-2001.
PD
XX 05-JUL-2000; 2000WO-US35727.
PF
XX 09-JUL-1999; 99US-0350841.
PR
XX (TRIM-) TRIMERIS INC.
PA
XX Jeffs P, Lackey JW, Erickson JB, Lawless MK, Merutka G;
PI
XX WPI; 2001-442157/47.
DR
XX Identifying a compound that inhibits the formation of or disrupts a
PT DP107/DP178 complex, especially compounds with antifusogenic, antiviral
PT or intracellular modulatory activity, by detecting the formation of a
```

```
PT DP107/DP178 complex -
XX
XX Disclosure; Fig 42; 259pp; English.
XX
XX The present invention relates to peptides which exhibit anti-retroviral
CC activity. The peptides of the invention (AAU12559-AAU14009) comprise
CC DP178-like and DP107-like peptides. The DP178 peptide corresponds
CC to amino acids 639-673 of the transmembrane protein gp41 from human
CC immunodeficiency virus 1 (HIV-1) isolate LAI. The DP107 peptide
CC corresponds to amino acids 538-595 of gp41 from HIV-1LAI. The invention
CC also relates to a method of identifying compounds that inhibit the
CC formation of or disrupts a DP107/DP178 complex. The method comprises
CC detecting the formation of a DP107/DP178 complex, both in the presence
CC or absence of a test compound, in a reaction mixture containing DP107
CC and DP178 peptides. The method is useful for identifying compounds,
CC including small molecule compounds, which may themselves exhibit
CC antifusogenic, antiviral or intracellular modulatory activity. The
CC DP178-like/DP107-like peptides are useful to inhibit human and non-human
CC retroviral, particularly HIV, transmission to uninfected cells. The
CC present sequence represents a peptide sequence from Staphylococcus aureus
CC enterotoxin A.
XX
SQ Sequence 257 AA;
Query Match 77.8%; Score 948; DB 22; Length 257;
Best Local Similarity 76.4%; Pred. No. 4.6e-88;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;
Qy 1 SEKSEINEKDLRKSELGQTALGNLKOIYYNYSKAITSSSEKSAQDFTLTLFKGFFTG 60
Db 25 SEKSEINEKDLRKSELGQTALGNLKOIYYNYSKAITSSSEKSAQDFTLTLFKGFFTG 84
Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 85 HSWYNDLLVDLFGSKDIDVYKGGKVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 144
Qy 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
Db 145 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
Qy 181 RGLIVFHSSEGSTVSYDLFDAAGQYSPDFTLLRIYRDNNTTISSTLSLSISLYLTT 233
Db 205 RGLIVFHTSTEPSVNYDLFDAAGQYSNTLLRIYRDNKNTINSENMHIDIYLYTS 257
RESULT 10
AAW06738
ID AAW06738 standard; Protein; 233 AA.
XX
XX AAW06738;
XX
XX 08-MAR-1997 (first entry)
DT
DE Staphylococcus enterotoxin A.
XX
XX Enterotoxin A; superantigen; antigen; cytokine; chemokine; T cell;
KW lymphocyte; monocyte; natural killer cell; gene therapy; cancer;
KW vaccine; adjuvant.
XX
OS Staphylococcus sp.
XX
XX WO9636366-A1.
PN
XX 21-NOV-1996.
PD
XX 20-MAY-1996; 96WO-US07432.
PF
XX 29-DEC-1995; 95US-0580806.
PR
XX 18-MAY-1995; 95US-0446918.
XX
XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY MED.
PA
XX Dow SW, Elmslie RE, Potter TA;
PI
```


XX WPI: 1997-011857/01.
 DR N-PSDB; AAT45699.
 XX
 PT Recombinant molecule encoding superantigen and opt. cytokine or
 PT chemokine - controls activity of effector cells (T cells, monocytes,
 PT natural killer cells), used for gene therapy of cancer
 XX
 XX Example 1; Page 98-99; 131pp; English.
 PS
 XX A cDNA clone (AAT45699) codes for staphylococcal enterotoxin A
 CC (AW06738) superantigen. . Nucleic acids encoding superantigens (see
 CC also AW06737, AAW06739), esp. truncated forms of the superantigen
 CC lacking the leader peptide, can be used in the gene therapy of
 CC cancer, infectious diseases and immunological disorders. The
 CC nucleic acid, optionally in combination with cytokine or chemokine
 CC nucleic acids, is delivered to an animal using e.g. liposomes. It
 CC acts by controlling the activity of effector cells, such as T-cells,
 CC macrophages, monocytes and/or natural killer cells. Localised
 CC prodn. of an effective but non-toxic amount of encoded proteins
 CC allows safe treatment of the animal.
 XX
 XX Sequence 233 AA;

Query Match 77.5%; Score 944; DB 18; Length 233;
 Best Local Similarity 76.3%; Pred. No. 1e-87;
 Matches 177; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy. 2 EKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSKSAQOFLTNTLLFKGFFTG 61
 Db 2 EKSEINEKDLRKSELOGTALGNLKOIYYNNEKAKTENKESHQFLQHTLFLKGFDDH 61
 Qy 62 PWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLTE 121
 Db 62 SWYNDLLVDFDSKDIDVYKGGKVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLTE 121
 Qy 122 EKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDQARHYLHGKFGLYNSDSFGGKQV 181
 Db 122 EKKVPINLWIDGKQNTVPLETNKKNTVQELDQARRYLQEKYLYNSDVFDDGKQV 181
 Qy 182 GLVPHSSEGSTVSYDLFDAQGOYPTLLRIYRNTTISSTLSLSLYLYTT 233
 Db 182 GLVPHSTSTPSVNDLFGAQQYNTLLRIYRDNKTINSENMHDIYLYTS 233

RESULT 11
 AAR45011
 ID AAR45011 standard; protein; 233 AA.
 XX
 AC AAR45011;
 XX
 DT 08-JUN-1994 (first entry)
 XX
 DE Staphylococcal enterotoxin SEA.
 XX
 KW Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
 KW autoimmune disease; toxicity; Protein A; perfusion system.
 XX
 OS Staphylococcus aureus.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 49
 FT /note= "Given in the specification as O, no further
 FT details given"
 XX
 PN WO9324136-A.
 XX
 PD 09-DEC-1993.
 XX
 XX 01-JUN-1993; 93WO-US05213.
 PF
 XX 01-JUN-1992; 92US-0891718.
 PR
 XX

PA (STON/) STONE J L.
 PA (TERM/) TERMAN D S.
 XX
 PI Stone JL, Terman DS;
 XX
 XX WPI: 1993-405418/50.
 DR
 XX
 PT Use of staphylococcal enterotoxin(s) and homologues - for
 PT treating cancer in a patient or for the treatment of auto-immune
 PT diseases
 XX
 XX Disclosure; Fig 1; 90pp; English.
 PS
 XX The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
 CC which may be used in the methods of the invention for treating cancer
 CC in a patient. These SEs, and homologues of them, can be used as
 CC tumouricidal agents for treating cancers and autoimmune disease.
 CC They exhibit tumouricidal activity and toxicity identical to that
 CC observed for the Protein A perfusion system. They may be administered
 CC by i.v. injection.
 XX
 XX Sequence 233 AA;

Query Match 77.3%; Score 942; DB 14; Length 233;
 Best Local Similarity 76.0%; Pred. No. 1.6e-87;
 Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNNSKAITSSKSAQOFLTNTLLFKGFFTG 60
 Db 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNNEKAKTENKESHQFLXHTLFLKGFETD 60
 Qy 61 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT 120
 Db 61 HSWYNDLLVDFDSKDIDVYKGGKVDLYGAYGYOCAGTGNKTKACMYGGVTLHDNNRLT 120
 Qy 121 EKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDQARHYLHGKFGLYNSDSFGGKQV 180
 Db 121 EKKVPINLWIDGKQNTVPLETNKKNTVQELDQARRYLQEKYLYNSDVFDDGKQV 180
 Qy 181 RGLVPHSSEGSTVSYDLFDAQGOYPTLLRIYRNTTISSTLSLSLYLYTT 233
 Db 181 RGLVPHSTSTPSVNDLFGAQQYNTLLRIYRDNKTINSENMHDIYLYTS 233

RESULT 12
 AAW35373
 ID AAW35373 standard; peptide; 233 AA.
 XX
 AC AAW35373;
 XX
 DT 20-APR-1998 (first entry)
 XX
 DE Staphylococcus enterotoxin SEA wild-type superantigen.
 XX
 KW SEE; SEA; Staphylococcus enterotoxin; superantigen; conjugate;
 KW treatment; cancer; infection; autoimmune disease; antibody.
 XX
 OS Staphylococcus sp.
 XX
 PN WO9736932-A1.
 XX
 PD 09-OCT-1997.
 XX
 PF 26-MAR-1997; 97WO-SE00537.
 XX
 PR 12-AUG-1996; 96US-0695692.
 PR 29-MAR-1996; 96SE-0001245.
 XX
 XX (PHAA) PHARMACIA & UPJOHN AB.
 PA
 XX Abrahmsen L, Antonsson P, Bjoerk P, Dohlsten M;
 PI Forsberg G, Hansson J, Kalland T;
 XX

DR WPI; 1997-503052/46.
 XX Conjugate of target seeking moiety and modified superantigen -
 PT useful for activating the immune system to treat cancer, viral
 PT infections, parasitic infestations and autoimmune diseases
 XX
 XX Claim 8; Pages 36-37; 58pp; English.
 PS
 PS This is the wild-type Staphylococcus enterotoxin SEA superantigen. This
 CC SEA superantigen can be modified to be used in a novel conjugate. The
 CC novel conjugate comprises a target seeking moiety and a modified wild
 CC type superantigen. The modified superantigen retains its ability to
 CC activate a subset of T cells, even though 1 or more wild-type amino acid
 CC residues in at least 1 region which functions in determining binding to
 CC T cell receptor (TCR) and activation of a subset of T cells has/have been
 CC replaced. Such a modified superantigen can optionally be used as part of
 CC a conjugate with a target seeking moiety, for activating the immune
 CC system to treat a mammalian disease. A pharmaceutical composition can be
 CC prepared comprising a modified antibody (preferably a Fab fragment fused
 CC to a peptide moiety providing activation of T cells in Vbeta specific
 CC manner) in which cysteines providing for interchain cysteine linkages in
 CC the native antibody have been replaced (preferably by serine residues) to
 CC prohibit cysteine formation. The modified wild-type superantigen is used
 CC for treating cancer, viral infections, parasitic infestations and
 CC autoimmune disease. The modified wild type superantigen has a lower
 CC immunogenicity and reactivity with neutralising antibodies and has fewer
 CC side-effects when used as a drug, compared to wild type superantigen.
 XX
 SQ Sequence 233 AA;

Query Match 77.3%; Score 941; DB 18; Length 233;
 Best Local Similarity 76.0%; Pred. No. 2.le-87;
 Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;
 Qy 1 SEKSEENEDLRKSELOGTALGNLQIYYNSKAITTSSEKSADQFLNTLLFKGFTG 60
 Db 1 SEKSEENEDLRKSELOGTALGNLQIYYNEKAKTENKESHQDLQHTLILFKGFTD 60
 Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
 Db 61 HSWYNDLLVDLFDSDKIDVKYKGVLDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
 Qy 121 EEKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLGYNDSFGKQV 180
 Db 121 EEKVPINLWIDGKQNTVPLETATKNTKNTVQBELDLQARRYLQEKYNLNSDVFDEKQV 180
 Qy 181 RGLIVFHSSEGSTVSYDLFDAQGOYPTLLRIYRDNTTISLSISLYLTT 233
 Db 181 RGLIVFHTSTEPSVNYDLFGAQQOYNTLLRIYRDNTINSENHIDIYLYTS 233

RESULT 13
 AAB67338
 ID AAB67338 standard; peptide; 233 AA.
 XX
 AC AAB67338;
 XX
 DT 23-APR-2001 (first entry)
 XX
 DE Staphylococcus aureus enterotoxin A protein.
 XX
 KW Tumour; cancer; immune; enterotoxin.
 XX
 OS Staphylococcus aureus.
 XX
 PN US6180097-B1.
 XX
 PD 30-JAN-2001.
 XX
 PF 30-OCT-1998; 98US-0183437.
 XX
 XX 31-JAN-1994; 94US-0189424.
 PR 19-JUN-1995; 95US-0491746.
 PR

PR 03-OCT-1989; 89US-0416530.
 PR 17-JAN-1990; 90US-0466577.
 PR 17-JAN-1991; 91WO-US00342.
 PR 01-JUN-1992; 92US-0891718.
 PR 02-MAR-1993; 93US-0025144.
 XX (TERM/) Terman D S.
 PA Terman DS;
 XX
 PI WPI; 2001-1586657/16.
 DR
 XX
 XX Tumor cell capable of stimulating antitumor immune reactivity in vitro
 PT or in vivo comprises exogenous nucleic acids encoding a superantigen
 PT and a costimulatory molecule
 XX
 XX Disclosure; Fig 2; 16pp; English.
 XX
 CC The present invention relates to a tumour cell capable of stimulating
 CC antitumor immune reactivity in vitro or in vivo contains and
 CC expresses an exogenous nucleic acid molecule encoding a superantigen
 CC or its active fragment and an exogenous nucleic acid molecule
 CC encoding a costimulatory molecule that activates T cells in
 CC conjunction with an antigenic stimulus. The invention may be used
 CC for cancer therapy by stimulating an anticancer immune response
 CC in vivo or ex vivo.
 XX
 SQ Sequence 233 AA;

Query Match 77.3%; Score 941; DB 22; Length 233;
 Best Local Similarity 76.0%; Pred. No. 2.le-87;
 Matches 177; Conservative 21; Mismatches 35; Indels 0; Gaps 0;
 Qy 1 SEKSEENEDLRKSELOGTALGNLQIYYNSKAITTSSEKSADQFLNTLLFKGFTG 60
 Db 1 SEKSEENEDLRKSELOGTALGNLQIYYNEKAKTENKESHQDLQHTLILFKGFTD 60
 Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
 Db 61 HSWYNDLLVDLFDSDKIDVKYKGVLDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
 Qy 121 EEKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLGYNDSFGKQV 180
 Db 121 EEKVPINLWIDGKQNTVPLETATKNTKNTVQBELDLQARRYLQEKYNLNSDVFDEKQV 180
 Qy 181 RGLIVFHSSEGSTVSYDLFDAQGOYPTLLRIYRDNTTISLSISLYLTT 233
 Db 181 RGLIVFHTSTEPSVNYDLFGAQQOYNTLLRIYRDNTINSENHIDIYLYTS 233

RESULT 14
 ABB76234
 ID ABB76234 standard; Protein; 233 AA.
 XX
 AC ABB76234;
 XX
 DT 09-AUG-2002 (first entry)
 XX
 DE Staphylococcus aureus enterotoxin A.
 XX
 KW Enterotoxin A; SEA; superantigen; antigen; tumour; cancer;
 KW antitumour; therapy.
 XX
 OS Staphylococcus aureus.
 XX
 XX Key Location/Qualifiers
 FT Misc-difference 49
 FT /note= "amino acid residue given as 'O' in the
 FT specification"
 XX
 XX US2002051765-A1.
 PN
 XX 02-MAY-2002.
 PD

```
XX PF 19-DEC-2000; 2000US-0741503.
XX PR 31-JAN-1994; 94US-0189424.
XX PR 19-JUN-1995; 95US-0491746.
XX PR 03-OCT-1989; 89US-0416530.
XX PR 17-JAN-1990; 90US-0466577.
XX PR 17-JAN-1991; 91WO-US00342.
XX PR 01-JUN-1992; 92US-0891718.
XX PR 02-MAR-1993; 93US-0025144.
XX PA (TERM/) TERMAN D S.
XX PI Terman DS;
XX DR WPI: 2002-415198/44.
XX PT Reagent for treating cancer without the need for e.g. radiotherapy,
XX PT comprises a specific V beta subset of T cells sensitized to a growing
XX PT tumor and stimulated with superantigens -
XX PS Disclosure; Fig 2; 17pp; English.
XX CC The present sequence is the protein sequence of enterotoxin A
XX CC (SEA) of Staphylococcus aureus. Similarity is shown, in several
XX CC stretches of sequence, between staphylococcal enterotoxins,
XX CC streptococcal pyrogenic exotoxins and staphylococcal exfoliative
XX CC toxins (see ABB76234-44). In the present invention, synthetic
XX CC polypeptides useful in tumour therapy and in blocking or destroying
XX CC autoreactive T and B lymphocyte populations are characterised by
XX CC substantial structural homology to staphylococcal enterotoxin A and
XX CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with
XX CC statistically significant sequence homology and similarity (Z value
XX CC of Lipman and Pearson algorithm in Monte Carlo analysis exceeding
XX CC 6) to include alignment of cysteine residues and similar hydropathy
XX CC profiles. These superantigens are used to treat solid tumours,
XX CC including their metastases, without radiation, surgery or
XX CC standard chemotherapeutic agents. A claimed method of human cancer
XX CC treatment involves contacting haematopoietic cells from a patient
XX CC with one or more superantigens ex vivo to generate stimulated cells,
XX CC selecting a specific V beta subset of cells, and reintroducing
XX CC these cells into the patient to induce an in vivo therapeutic,
XX CC tumouricidal reaction.
XX SQ Sequence 233 AA;
    Query Match 76.6%; Score 933; DB 23; Length 233;
    Best Local Similarity 75.5%; Pred. No. 1.3e-86;
    Matches 176; Conservative 21; Mismatches 36; Indels 0; Gaps 0;
QY 1 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITSSEKSADQFLNTLLFGFFTG 60
Db 1 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITENKESHDOFLXHTILFKGFFTD 60
QY 61 HPWYNDLLVDLGSTATSEYSGSSVDLYGAYCYQCAGGTGPNKTACMGVTLHDNRLLT 120
Db 61 HSWYNDLLVDLDFSDKIDVRYKGGKVDLYGAYCYQCAGGTGPNKTACMGVTLHDNRLLT 120
QY 121 EEKKVPINLWIDGQKQTPVIDKVKTSKEVTVQELDLQARHVLHGKFGLYNSDFGKQV 180
Db 121 EEKKVPINLWLDGQKQNTPLETVTKTKNNKNTVQELDPQARRYLQEKYNYNSDFGKQV 180
QY 181 RGLIVPHSSEGSFVSYDLFDAQOQPTDLLRIYRONTTISSTSLSLSLYLYTT 233
Db 181 RGLIVFTSTEPSVNYDLFGAGQISNTLLRIYRDNKTINSENHIDIYLYTS 233
RESULT 15
ABB79501
XX ID ABB79501 standard; Protein; 257 AA.
XX AC ABB79501;
XX DB 25 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITENKESHDOFLXHTILFKGFFTD 84
```

```
DT 23-SEP-2002 (first entry)
XX Staphylococcal enterotoxin A vaccine, periplasmic (A489270P).
XX DE Enterotoxin A; superantigen; antigen; toxin; vaccine; A489270P;
XX KW attenuation; mutant; mutein.
XX OS Staphylococcus sp.
XX OS Synthetic.
XX FH Key Location/Qualifiers
FT Peptide 1..25
FT Protein /label= signal_peptide
FT /label= Mature_protein
FT Misc-difference 72
FT /note= "wild-type Leu substituted by Arg"
FT Misc-difference 94
FT /note= "wild-type Asp substituted by Arg"
FT Misc-difference 116
FT /note= "wild-type Tyr substituted by Ala"
XX US6399332-B1.
XX 04-JUN-2002.
XX 01-SEP-1998; 98US-0144776.
XX 25-JUN-1997; 97US-0882431.
XX (USSA ) US SEC OF ARMY.
XX Ulrich RG, Olson MA, Bavari S;
XX WPI: 2002-546281/58.
XX N-PSDB; ABB84222.
XX Novel isolated and purified superantigen toxin DNA fragment which has
XX been genetically altered, useful for producing vaccine for treatment of
XX superantigen toxin-associated bacterial diseases -
XX Claim 4; Column 33-35; 46pp; English.
XX The present sequence is the protein sequence of staphylococcal
XX enterotoxin A (SEA) vaccine, periplasmic (A489270P). The vaccine
XX comprises 3 amino acid substitutions introduced into the SEA
XX sequence: L48R, Y89A and D70R. These mutations reduce the binding
XX of the toxin to major histocompatibility complex (MHC) Class II
XX and/or T cell receptors. The full-length expressed product is
XX secreted into the periplasmic space of Escherichia coli host cells,
XX and the leader peptide is recognised and cleaved by a native
XX mechanism. The vaccine is used to protect against superantigen
XX toxin infections. Superantigen attributes are absent, but the
XX superantigen is effectively recognised by the immune system and an
XX appropriate antibody response is produced. In examples from the
XX invention, attenuated superantigen toxins were shown to protect
XX animals against challenge with wild-type toxin. Methods of
XX producing and using the altered superantigen toxins as vaccines,
XX and in diagnosis and therapy, are provided. A multivalent
XX vaccine consisting of altered superantigen toxins from SEA, SEB,
XX SEC-1, TSET-1 and streptococcal SPEA is predicted to provide
XX protective immunity against the majority of bacterial superantigen
XX toxins.
XX Sequence 257 AA;
    Query Match 75.9%; Score 925; DB 23; Length 257;
    Best Local Similarity 75.1%; Pred. No. 1e-85;
    Matches 175; Conservative 21; Mismatches 37; Indels 0; Gaps 0;
QY 1 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITSSEKSADQFLNTLLFGFFTG 60
Db 25 SEKSEINEKDLRKSELOQTALGNLKIYYNNKRAITENKESHDOFLXHTILFKGFFTD 84
```


FT Misc-difference 2 /note= "Encoded by AG"
FT Misc-difference 18 /note= "Encoded by TTG"
FT Misc-difference 48 /note= "Wild type Leu substituted with Arg"
FT Misc-difference 70 /note= "Wild type Asp substituted with Arg"
FT Misc-difference 92 /note= "Wild type Tyr substituted with Ala"
FT Misc-difference 157 /note= "Encoded by CTT"
FT Misc-difference 180 /note= "Encoded by CAG"
XX
PN WO200009154-A1.
XX
PD 24-FEB-2000.
XX
XX 13-AUG-1998; 98WO-US16766.
XX
PR 13-AUG-1998; 98WO-US16766.
XX (REED-) REED ARMY INST RES WALTER.
XX
XX Ulrich RG, Olson MA, Bavari S;
PI
DR WPI: 2000-224177/19.
DR N-PSDB; AAZ51106.
XX
XX Nucleic acid encoding superantigen toxin useful as a vaccine and for
PT diagnosis of superantigen-associated bacterial infections -
XX
XX Claim 8; Page 74-76; 118pp; English.
XX
CC The present amino acid sequence is the mutant Staphylococcal enterotoxin
CC A (SEA), a bacterial superantigen toxin (SAG), used for the
CC formulation of SEA vaccine A489270P. The coding region of this SAG toxin
CC is altered by site directed mutagenesis, that results in disruption of
CC binding of the toxin to both the MHC class II or T-cell antigen
CC receptor. This altered SAG toxin has the leader peptide cleaved by native
CC bacterial enzymatic mechanism and the first residue of the mature protein
CC is encoded by the transcriptional start site (ATG). SEA has
CC antibacterial and cytostatic activity. This sequence is useful for the
CC production of SEA vaccines and specific antibodies. This vaccine
CC overcomes the disadvantages of the chemically inactivated toxoids and is
CC designed to protect individuals against one or several related
CC staphylococcal and streptococcal toxins. It is used for the diagnosis and
CC treatment or amelioration of superantigen-associated bacterial
CC infections.
XX
SQ Sequence 233 AA:
Query Match 74.1%; Score 903; DB 21; Length 233;
Best Local Similarity 73.7%; Pred. No. 1.5e-83;
Matches 171; Conservative 21; Mismatches 40; Indels 0; Gaps 0;
QY 2 EKSEETNEKDLRRKSELQGTALGNLKOIYYNSKATTSSEKSADOFLTNTLLFKGFFTG 61
DB 2 EKSEETNEKDLRRKSELQGTALGNLKOIYYNEKAKTENKESHDFRQITLFGFFTDH 61
QY 62 PYNOLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
DB 62 SWNDLLVRFDKSDIVDYKGGKVDLYGAYAGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
QY 122 EKKVPINLWIDGKQTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDFSGKVKOR 181
DB 122 EKKVPINLWIDGKQNTVPLETYTKNNVTVOELDKQARRYLQERYNLNSDFDGKVAR 181
QY 182 GLIVFHSSBGSVSYDLDFDAQGOYPTDLLRIYRDNNTTISSTLSLSLYLYTT 233
DB 182 GLIVFHTSTEPSVNDLDFGAQGOYNTLLRIYRDNKNTINSENNHIDIYLYTS 233

RESULT 20
AAR45013
XX AAR45013 standard; protein; 228 AA.
XX AAR45013;
XX 08-JUN-1994 (first entry)
XX Staphylococcal enterotoxin SED.
XX Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
KW autoimmune disease; toxicity; Protein A; perfusion system.
XX Staphylococcus aureus.
XX WO9324136-A.
XX 09-DEC-1993.
XX 01-JUN-1993; 93WO-US05213.
XX 01-JUN-1992; 92US-0891718.
XX (STON/) STONE J L.
XX (TERM/) TERMAN D S.
XX Stone JL, Terman DS;
XX WPI: 1993-405418/50.
XX Use of staphylococcal enterotoxin(s) and homologues - for
PT treating cancer in a patient or for the treatment of auto-immune
PT diseases
XX
XX Disclosure; Fig 1; 90pp; English.
XX The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
CC which may be used in the methods of the invention for treating cancer
CC in a patient. These SEs, and homologues of them, can be used as
CC tumouricidal agents for treating cancers and auto-immune disease.
CC They exhibit tumouricidal activity and toxicity identical to that
CC observed for the Protein A perfusion system. They may be administered
CC by i.v. injection.
XX
SQ Sequence 228 AA;
Query Match 49.7%; Score 605; DB 14; Length 228;
Best Local Similarity 52.0%; Pred. No. 3.3e-53;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;
QY 7 INEKDLRRKSELQGTALGNLKOIYYNSKATTSSEKSADOFLTNTLLPKGFFTGHPWYND 66
DB 2 VKEKELHKKSELSTALNNMKHSYADKNPITGENKSTGDOFLNTLLYKKFFTDLINFED 61
QY 67 LLVLDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVP 126
DB 62 LLINFSKEMAHQFKSKNDVVPYPIRYINCYGEIDRTACTYGGVTPHEGNLKERKKIP 121
QY 127 INLWIDGKQTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDFSGKVKORGLIVF 186
DB 122 INLWINGVQKESLDKVTQDKNNVTVOELDAQARRYLQKDLKLYNNDTLGGKIQRGKIEF 181
QY 187 HSSEGSTSYDLDFDAQGOYPTDLLRIYRDNNTTISSTLSLSLYLY 231
DB 182 DSSDGSKVSVDLDFVKGDFPEKQLRIYSDNKTLSLSTLHLDIYLY 226
RESULT 21
AAB67340
XX AAB67340 standard; peptide; 228 AA.
XX AAB67340;
AC

XX 23-APR-2001 (first entry)
XX Staphylococcus aureus enterotoxin D protein.
XX Tumour; cancer; immune; enterotoxin.
XX Staphylococcus aureus.
XX US6180097-B1.
XX 30-JAN-2001.
XX 30-OCT-1998; 98US-0183437.
XX 31-JAN-1994; 94US-0189424.
XX 19-JUN-1995; 95US-0491746.
XX 03-OCT-1989; 89US-0416530.
XX 17-JAN-1990; 90US-0466577.
XX 17-JAN-1991; 91WO-US00342.
XX 01-JUN-1992; 92US-0891718.
XX 02-MAR-1993; 93US-0025144.
XX (TERM/) TERMAN D S.
XX Terman DS;
XX WPI; 2001-158657/16.
XX Tumor cell capable of stimulating antitumor immune reactivity in vitro
PT or in vivo comprises exogenous nucleic acids encoding a superantigen
PT and a costimulatory molecule -
XX Disclosure; Fig 2; 16pp; English.
XX The present invention relates to a tumour cell capable of stimulating
CC antitumor immune reactivity in vitro or in vivo contains and
CC expresses an exogenous nucleic acid molecule encoding a superantigen
CC or its active fragment and an exogenous nucleic acid molecule
CC encoding a costimulatory molecule that activates T cells in
CC conjunction with an antigenic stimulus. The invention may be used
CC for cancer therapy by stimulating an anticancer immune response
CC in vivo or ex vivo.
XX Sequence 228 AA;
Query Match 49.7%; Score 605; DB 22; Length 228;
Best Local Similarity 52.0%; Pred. No. 3.3e-53;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;
Qy 7 INEKDLRKSELOGTALGNLKOIYYVNSKAITSSPKSADQFLTNTLLPKGFTGHPWYND 66
Db 2 VKERELHKKSELSSTALNMKHSYADKNPIIGENKSTGDFLENTLLYKFFTDLINEF 61
Qy 67 LLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHDNRRLTEKKVP 126
Db 62 LLINPNSKEMAQHFKSNVDVPIRYSINCYGGEIDRTACTYGGVTPHEGNKLERKKIP 121
Qy 127 INLWIDGKQTTVPIDKVKTSKKEVTQVELDQARHYLHGKFGLYNSDSFGKVGRLIVF 186
Db 122 INLWINGVQKEVSLDKVQTDKKNVTQVELDAQARRYLQKDLKLYNNDTLGGKIQRGKIEF 181
Qy 187 HSSEGSTSYDLFDAQGYPPDTLLRIYRDNTTISSTLSISLYL 231
Db 182 DSSDGSKYSYDLFDVKGDFPEKQLRIYSDNKTLSLHIDILY 226
RESULT 22
ID ABB76236
XX ABB76236 standard; Protein; 228 AA.
XX AC ABB76236;
XX

DT 09-AUG-2002 (first entry)
XX Staphylococcus aureus enterotoxin D.
XX Enterotoxin D; SED; superantigen; antigen; tumour; cancer;
KW antitumour; therapy.
XX Staphylococcus aureus.
XX OS
XX US2002051765-A1.
XX 02-MAY-2002.
XX 19-DEC-2000; 2000US-0741503.
XX 31-JAN-1994; 94US-0189424.
XX 19-JUN-1995; 95US-0491746.
XX 03-OCT-1989; 89US-0416530.
XX 17-JAN-1990; 90US-0466577.
XX 17-JAN-1991; 91WO-US00342.
XX 01-JUN-1992; 92US-0891718.
XX 02-MAR-1993; 93US-0025144.
XX (TERM/) TERMAN D S.
XX Terman DS;
XX WPI; 2002-415198/44.
XX Reagent for treating cancer without the need for e.g. radiotherapy,
PT comprises a specific V beta subset of T cells sensitized to a growing
PT tumor and stimulated with superantigens -
XX Disclosure; Fig 2; 17pp; English.
XX The present sequence is the protein sequence of enterotoxin D
CC (SED) of Staphylococcus aureus. Similarity is shown, in several
CC stretches of sequence, between staphylococcal enterotoxins,
CC streptococcal pyrogenic exotoxins and staphylococcal exfoliative
CC toxins (see ABB76234-44). In the present invention, synthetic
CC polypeptides useful in tumour therapy and in blocking or destroying
CC autoreactive T and B lymphocyte populations are characterized by
CC substantial structural homology to staphylococcal enterotoxin A and
CC enterotoxin B, and to streptococcal pyrogenic exotoxins, with
CC statistically significant sequence homology and similarity (Z value
CC of Lipman and Pearson algorithm in Monte Carlo analysis exceeding
CC 6) to include alignment of cysteine residues and similar hydrophathy
CC profiles. These superantigens are used to treat solid tumours,
CC including their metastases, without radiation, surgery or
CC standard chemotherapeutic agents. A claimed method of human cancer
CC treatment involves contacting haematopoietic cells from a patient
CC with one or more superantigens ex vivo to generate stimulated cells,
CC selecting a specific V beta subset of cells, and reintroducing
CC these cells into the patient to induce an in vivo therapeutic,
CC tumouricidal reaction.
XX Sequence 228 AA;
Query Match 49.7%; Score 605; DB 23; Length 228;
Best Local Similarity 52.0%; Pred. No. 3.3e-53;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;
Qy 7 INEKDLRKSELOGTALGNLKOIYYVNSKAITSSPKSADQFLTNTLLPKGFTGHPWYND 66
Db 2 VKERELHKKSELSSTALNMKHSYADKNPIIGENKSTGDFLENTLLYKFFTDLINEF 61
Qy 67 LLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHDNRRLTEKKVP 126
Db 62 LLINPNSKEMAQHFKSNVDVPIRYSINCYGGEIDRTACTYGGVTPHEGNKLERKKIP 121
Qy 127 INLWIDGKQTTVPIDKVKTSKKEVTQVELDQARHYLHGKFGLYNSDSFGKVGRLIVF 186
Db 122 INLWINGVQKEVSLDKVQTDKKNVTQVELDAQARRYLQKDLKLYNNDTLGGKIQRGKIEF 181

XX Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
KW group A streptococcus; Streptococcus pyogenes; antibacterial;
KW antinflammatory; infection; vaccine; meningitis; gene therapy.
XX
OS Streptococcus pyogenes.
XX
PN WO200234771-A2.
XX
PD 02-MAY-2002.
XX
XX 29-OCT-2001; 2001WO-GB04789.
XX
XX 27-OCT-2000; 2000GB-0026333.
PR 24-NOV-2000; 2000GB-0028727.
PR 07-MAR-2001; 2001GB-0005640.
XX
XX (CHIR-) CHIRON SPA.
PA (GENO-) INST GENOMIC RES.
XX
XX Telford J, Masignani V, Margarit Ros YI, Grandi G, Fraser C;
PI Tettelin H;
XX
XX WPI: 2002-352536/38.
DR N-PSDB; ABN69988.
XX
XX New Streptococcus protein for the treatment or prevention of infection
PT or disease caused by Streptococcus bacteria, such as meningitis, and
PT for detecting a compound that binds to the protein -
XX
PS Claim 1; Page 3920; 4525pp; English.
XX
XX The invention relates to a protein (ABP25413-ABP30895) from group B
CC Streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
CC (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
CC the specification. The proteins have antibacterial and antiinflammatory
CC activity. (I), nucleic acids encoding (I), ABN6044-ABN71526 and
CC antibodies that bind (I) are used in the manufacture of medicaments for
CC the treatment or prevention of infection or disease caused by
CC Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
CC Nucleic acids encoding (I) are used to detect Streptococcus in a
CC biological sample. (I) is used to determine whether a compound binds to
CC used as a vaccine or diagnostic composition. The disease caused by
CC Streptococcus that is prevented or treated may be meningitis. Nucleic
CC acid encoding (I) may be used to recombinantly produce (I) and may be
CC used in gene therapy. Antibodies to (I) are used for affinity
CC chromatography, immunoassays, and distinguishing/identifying
CC Streptococcus proteins.
XX
SQ Sequence 259 AA;
Query Match 23.9%; Score 291.5; DB 23; Length 259;
Best Local Similarity 31.3%; Pred. No. 3.3e-21;
Matches 72; Conservative 42; Mismatches 91; Indels 25; Gaps 8;
QY 17 ELQGTALGNLKOIY-YNKSATTSSEKSADQFLTNLLFKGFTGHPWYN---DLLVDLG 72
DB 34 EMSVGVINLRNLSTYDPTVEVKGINSGPP-SCSLFYKNT----PYGNSIELKVELN 88
QY 73 SPAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEERKVPINLWID 132
DB 89 SVEKANFESGKRVDIFTLEYSPPCNISKNKNS---YGGITLSDGNRI-DKNKIPVNIIFD 144
QY 133 G---KQTTVPIDKVKTSKEVTVOELDQARHYLHGKGLY-----NDSDFGGKVQ 180
DB 145 GVOQKYSYTDISTGTDKKEVTIQELDKRSYRLOKHENFYGDVDFGSRSRQSGFE 204
QY 181 RGLIVFHSSEGSTVSYDLFDAQGQYPTLLRIYRDNNTTSSLSISLYL 230
DB 205 EGNIIIFHLSNGERISYNLFDTGCHGDRSMLKKYSNKTAYSQDLHIDIYL 254

RESULT 26
AAAY06254
ID AAY06254 standard; Protein; 239 AA.
XX
AC AAY06254;
XX
DT 23-AUG-1999 (first entry)
XX
DE Staphylococcal group C enterotoxin SEC3-FRI909.
KW Enterotoxin; SEC3-FRI909; toxin; disulfide loop;
KW protein engineering.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Disulfide-bond 93...110
XX
PN WO9927889-A2.
XX
PD 10-JUN-1999.
XX
PF 01-DEC-1998; 98WO-US25107.
XX
PR 02-DEC-1997; 97US-0067357.
XX
PA (IDAH-) IDAHO RES FOUND INC.
XX
PI Bohach GI;
XX
DR WPI: 1999-358008/30.
XX
XX Non-toxic modified staphylococcal enterotoxins
PS Disclosure; Page 17; 25pp; English.
XX
XX This protein represents the Staphylococcus aureus type C
CC enterotoxin SEC3-FRI909. The invention relates to pyrogenic toxins,
CC such as staphylococcal enterotoxins, modified in the disulfide loop
CC region. Typically, the modification involves deletions within the
CC disulfide loop region of SEC (see AAY06261). The modified toxins
CC retain useful biological properties, such as the ability to induce
CC cytokine production, but have substantially reduced toxicity
CC compared to the corresponding unmodified native toxin. Emetic
CC response inducing activity and fever inducing activity are
CC typically decreased by at least about 100-fold, while LD50 (in Dutch
CC Belted rabbits) is at least 100-fold higher than the native toxin.
XX
SQ Sequence 239 AA;
Query Match 23.6%; Score 287.5; DB 20; Length 239;
Best Local Similarity 33.3%; Pred. No. 7.5e-21;
Matches 78; Conservative 45; Mismatches 92; Indels 19; Gaps 9;
QY 11 DLRRKSELQGTALGNLKOIYYNKSATTSSE-KSADQFLTNLLFKGFTGHPWYNDLLV 69
DB 10 DLHKSSSETGT-MGNMK--YLYDDHYVSATKVKSVYDKFLAHLIYNINDKLNNDKYKT 56
QY 70 DLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGGVTLHDNNRLTEE 122
DB 67 ELLNEDLANKYKDEVDVYGSNNYVNCYFSSKDNVGVTSKGTCTMYGGITKHEGNHFDNG 126
QY 123 --KKYPINLWIDGKQTTVPIDKVKTSKEVTVOELDQARHYLHGKGLYNSDSFGKVQ 180
DB 127 NLQNLIRYV-ENKRNITISFE-VQTDKKSVTAAQELDIKARNFLINKNLNLYEFSN--SPYE 182
QY 181 RGLIVFHSSEGSTVSYDLFDAQGQYPT--TLIRIYRDNNTTSSLSISLYLT 232
DB 183 TCYIKFIESNGNTFWYDMMPAGDKFDQSKYLMYIYDKNMVDSKSVKIEVHLTT 236
RESULT 27
AAAY06253

DT 23-AUG-1999 (first entry)
XX Staphylococcal group C enterotoxin SEC-MNCopeland.
DE Enterotoxin; SEC-MNCopeland; toxin; disulfide loop;
XX protein engineering.
KW Staphylococcus aureus.
XX

XX Key Location/Qualifiers
FH Disulfide-bond 93..110
XX WO9927889-A2.
XX 10-JUN-1999.
XX

XX 01-DEC-1998; 98WO-US25107.
XX 02-DEC-1997; 97US-0067357.
XX (IDAH-) IDAHO RES FOUND INC.
XX Bohach GI;
XX WPI; 1999-358008/30.
XX

XX Non-toxic modified staphylococcal enterotoxins
XX Disclosure; Page 17; 25pp; English.
XX This protein represents the Staphylococcus aureus type C
CC enterotoxin SEC-MNCopeland. The invention relates to pyrogenic
CC toxins, such as staphylococcal enterotoxins, modified in the
CC disulfide loop region. Typically, the modification involves
CC deletions within the disulfide loop region of SEC (see AAY06261).
CC The modified toxins retain useful biological properties, such as
CC the ability to induce cytokine production, but have substantially
CC reduced toxicity compared to the corresponding unmodified native
CC toxin. Emetic response inducing activity and fever inducing
CC activity are typically decreased by at least about 100-fold, while
CC LD50 (in Dutch Belted rabbits) is at least 100-fold higher than
CC the native toxin.
XX

SQ Sequence 239 AA;
Query Match 23.0%; Score 280.5; DB 20; Length 239;
Best Local Similarity 31.3%; Pred. No. 3.8e-20;
Matches 76; Conservative 46; Mismatches 102; Indels 19; Gaps 9;
Qy 2 EKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSE-KSADQFLNTLLFKGFTG 60
Db 1 ESQDPPTDELHKSEFTGT-MGNMK--YLYDDHYVSATKVKSVDFLAHDLIYNISDKK 57
Qy 61 HPWYNDLLVGLGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACHYGGVTL 113
Db 58 LKNYDKVKTTELLNEDLAKKYDEVVDVYGSNYVNCYFSSKDNVGVGKTCMYGGITK 117
Qy 114 HDNNRLTEE--KKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYN 171
Db 118 HEGNHFNDNGNLQNLIRVY-ENKRTISFE-VQTDKKSVAQELDIKARNLINKKNLYE 175
Qy 172 SDSFGGKVGORGLIVFHSSEGTSVSYDLFAQGVDP--TLLRIYRDNTTISSTLSISLY 229
Db 176 FNS--SPVETGYIKFIENNGNTFOYDMMPAPGDKFQDSKYLMMYNDKNTVDSKRVKIEVH 233
Qy 230 LYT 232
Db 234 LTT 236

RESULT 30
ID AAY06252 standard; Protein; 239 AA.

XX AAY06252;
XX 23-AUG-1999 (first entry)
DT Staphylococcal group C enterotoxin SEC2.
XX Enterotoxin; SEC2; toxin; disulfide loop; protein engineering.
XX Staphylococcus aureus.
XX

XX Key Location/Qualifiers
FH Disulfide-bond 93..110
XX WO9927889-A2.
XX 10-JUN-1999.
XX 01-DEC-1998; 98WO-US25107.
XX 02-DEC-1997; 97US-0067357.
XX (IDAH-) IDAHO RES FOUND INC.
XX Bohach GI;
XX WPI; 1999-358008/30.
XX

XX Non-toxic modified staphylococcal enterotoxins
XX Disclosure; Page 17; 25pp; English.
XX This protein represents the Staphylococcus aureus type C
CC enterotoxin SEC2. The invention relates to pyrogenic toxins, such
CC as staphylococcal enterotoxins, modified in the disulfide loop
CC region. Typically, the modification involves deletions within the
CC disulfide loop region of SEC (see AAY06261). The modified toxins
CC retain useful biological properties, such as the ability to induce
CC cytokine production, but have substantially reduced toxicity
CC compared to the corresponding unmodified native toxin. Emetic
CC response inducing activity and fever inducing activity are
CC typically decreased by at least about 100-fold, while LD50 (in Dutch
CC Belted rabbits) is at least 100-fold higher than the native toxin.
XX

SQ Sequence 239 AA;
Query Match 22.9%; Score 279.5; DB 20; Length 239;
Best Local Similarity 31.3%; Pred. No. 4.9e-20;
Matches 76; Conservative 46; Mismatches 102; Indels 19; Gaps 9;
Qy 2 EKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSE-KSADQFLNTLLFKGFTG 60
Db 1 ESQDPPTDELHKSEFTGT-MGNMK--YLYDDHYVSATKVKSVDFLAHDLIYNISDKK 57
Qy 61 HPWYNDLLVGLGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACHYGGVTL 113
Db 58 LKNYDKVKTTELLNEDLAKKYDEVVDVYGSNYVNCYFSSKDNVGVGKTCMYGGITK 117
Qy 114 HDNNRLTEE--KKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYN 171
Db 118 HEGNHFNDNGNLQNLIRVY-ENKRTISFE-VQTDKKSVAQELDIKARNLINKKNLYE 175
Qy 172 SDSFGGKVGORGLIVFHSSEGTSVSYDLFAQGVDP--TLLRIYRDNTTISSTLSISLY 229
Db 176 FNS--SPVETGYIKFIENNGNTFOYDMMPAPGDKFQDSKYLMMYNDKNTVDSKRVKIEVH 233
Qy 230 LYT 232
Db 234 LTT 236

RESULT 31
AAR45016

Db	126	NLQNLVLRVY-ENKRNWISFE-VQTDKKSVTAQELDIKARNFLINKKNLYEENS--SPYE 181
Qy	181	RLGIVFHSSGSGSVSYDLFAQGO--YPDTLLRIYRNTTISSTLSISLYLT 232
Db	182	TGYIKFIENNQFWYDLMPAPGDKFKQSKYLMYNDNKVTDSKSVKIEVHLT 235
RESULT 34		
AAAG63856		standard; Protein; 228 AA.
XX		
AC	AAAG63856;	
XX		
DT	29-OCT-2001 (first entry)	
XX		
DE	Amino acid sequence of a modified Staphylococcal enterotoxin C1.	
XX		
KW	Enterotoxin C1; SEC-SER; infectious disease; mastitis.	
OS	Synthetic.	
OS	Staphylococcus sp.	
XX		
PN	W0200160851-A1.	
XX		
PD	23-AUG-2001.	
XX		
PF	31-OCT-2000; 2000WO-KR01241.	
XX		
PR	17-FEB-2000; 2000KR-0007612.	
XX		
PA	(GLDS) LG CHEM LTD.	
XX		
PI	Lee H, Park Y, Han K, Chang B, Lee Y;	
XX		
DR	WPI; 2001-522585/57.	
DR	N-PSDB; AAH74983.	
XX		
PT	Producing stable modified Staphylococcal toxin polypeptide for treating infectious diseases, e.g. mastitis, in animals, involves substituting cysteine at specified position of modified Staphylococcal toxin C1 with serine -	
XX		
PS	Claim 1; Page 59-60; 64pp; English.	
XX		
CC	The present sequence represents a modified Staphylococcal enterotoxin designated SEC-SER. The modified toxin is characterised in that the 9th amino acid (cysteine) is substituted with serine. This results in a toxin that has improved stability. The modified enterotoxin is useful for preventing or treating infectious diseases due to microorganisms in animals such as cows, pigs, horses, sheep, hens, dogs and cats, e.g. mastitis in cows.	
XX		
SQ	Sequence 228 AA;	
Query Match	22.5%;	Score 273.5; DB 22; Length 228;
Best Local Similarity	32.1%;	Pred. No. 1.9e-19;
Matches	76; Conservative 45; Mismatches 97; Indels 19; Gaps	
Qy	2	EKSEINEKDLRKSELOCTALGNLKOIY--YNNSKAITSEKSADQFLNTLLFKGFFT 59
Db	2	ESQPDPTPDELHKASKFTG-ILMENKVLVDHYVS--ATKVKSVDRFLAHLDTIYNISDK 57
Qy	60	GHPWYNLLVDLGGSTAATSEYEGSSVDLYGAYCYOCAGGTPNKTACMGVTLHDNRL 119
Db	58	KLWYDKVKTELLNEGLAKYTKDEVDVYSGNYVNCSGKT-----CMYGITKHGSHNF 112
Qy	120	TEE--KKVPINLWDGKQTTPIDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGG 177
Db	113	DNGNLQNLVLRVY-ENKRNWISFE-VQTDKKSVTAQELDIKARNFLINKKNLYEENS--S 160
Qy	178	KVQRGIVFHSSGSGSVSYDLFAQGOYD--TLRIYRNTTISSTLSISLYLT 232
Db	169	PYETGYIKFIENNQFWYDLMPAPGDKFKQSKYLMYNDNKVTDSKSVKIEVHLT 225

```
RESULT 35
AAR13208
ID AAR13208 standard; Protein; 238 AA.
XX
AC AAR13208;
XX
DT 15-OCT-1991 (first entry)
XX
DE Staphylococcal enterotoxin C3.
XX
KW SEC3; cancer treatment; pyrogen; tumouricide.
XX
OS Staphylococcus aureus.
XX
PN WO9110680-A.
XX
PD 25-JUL-1991.
XX
PF 17-JAN-1991; 91WO-US00342.
XX
PR 17-JAN-1990; 90US-0466577.
XX
PA (TERM/) Terman D S.
XX
PI Terman DS;
XX
DR WPI; 1991-237984/32.
XX
PT Treating cancer with enterotoxin from Staphylococcus aureus -
PT administered by IV injection, having same tumoricidal activity
PT as Staphylococcal protein A without potential toxic reactions
XX
PS Disclosure; Fig 1; 74pp; English.
XX
CC SEC3 was isolated and purified from S.aureus. It can be used for
CC treating cancer, activating cytokine mediators and procoagulant
CC systems, augmenting natural killer cell cytotoxicity, etc. The
CC enterotoxin can be administered intravenously, optionally with
CC ibuprofen to attenuate toxic reaction to SEC3. Synthetic
CC polypeptides having structural homology to Staphylococcal
CC are claimed, provided the homology includes statistically
CC significant sequence homology, alignment of Cysteine residues and
CC similar hydrophathy profiles.
CC See AAR13203-R13211.
XX
SQ Sequence 238 AA;
Query Match 22.4%; Score 272.5; DB 12; Length 238;
Best Local Similarity 31.1%; Pred. No. 2.5e-19;
Matches 76; Conservative 41; Mismatches 88; Indels 39; Gaps 8;
Qy 11 DLRRKSELQGTALGNLQIYYNSKAITSSSE-KSADQFLTNTLLFKGFTGHPWYNLLLV 69
Db 9 ELHKSSEFTGT-MGNMK--YLYDDHYVSATKVMYSVDKELAHDLIYNSDKKLKNDYKVT 65
Qy 70 DLGSTAATSEYEGSSVDLYGAYGQC-----AGGTPNKATCMYGGVTLHDNN----- 117
Db 66 ELLNEDLAKYKDEVDVYGSNNYNYCFSSKDNNGKVTGKTCMYGGITRHEGNHGDNG 125
Qy 118 -----RLTEKKVPINLWIDGKQTTVPIDKVTSSKEVTVQELDLQARHYLHGKFGLY 170
Db 126 NLQNLIRVYENKRWTSF-----EVQTDKSVTAQELDIKARNFLINKKNLY 173
Qy 171 NSDSFGGKVGRLIYVHSEGSTSVSYDLDAQOQYPD--TLRIYRDNTTISSTLSISL 228
Db 174 EPNs--SPYETGYINFIENNGTWFYDLMPAPGDFDQSLYLMYNDKNKTVDKSVKIEV 231
Qy 229 YLYT 232
Db 232 HLTT 235
```

```
RESULT 36
AAW12153
ID AAW12153 standard; Protein; 251 AA.
XX
AC AAW12153;
XX
DT 04-NOV-1997 (first entry)
XX
DE Streptococcus pyogenes Streptococcal toxin A mutant Cys90Ser.
XX
KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
KW vaccine; protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
KW fever; hypotension; group A streptococcal infection; myositis;
KW fascitis; liver damage; T cell; lymphoma; ovarian; uterine.
XX
OS Streptococcus pyogenes.
XX
PD Synthetic.
XX
PF Key Location/Qualifiers
FT Peptide 1..30
FT Peptide /label= sig_peptide
FT Peptide 31..251
FT Misc-difference 120 /label= mat_peptide
FT /note= "wild type Cys replaced by Ser"
XX
PN WO9640930-A1.
XX
PD 19-DEC-1996.
XX
PR 07-JUN-1996; 96WO-US10252.
XX
PR 07-JUN-1995; 95US-0480261.
XX
PA (MINU ) UNIV MINNESOTA.
XX
PI Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
XX
DR WPI; 1997-099936/09.
XX
PT Mutant SPE-A toxin with at least one amino acid change is
PT substantially non-lethal - used in vaccine composition for
PT treatment of cancer and streptococcal toxic shock syndrome etc.
XX
PS Example 4; Page -: 102pp; English.
XX
CC The present sequence is a non-lethal Streptococcus pyogenes
CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce
CC vaccines to protect animals against wild type SPE-A and to treat
CC cancer and streptococcal toxic shock syndrome (STSS). The mutant
CC SPE-A causes neutralising antibodies (Ab) to be produced, which
CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,
CC group A streptococcal infection, myositis, fascitis and liver
CC damage. The neutralising Ab is preferably administered in
CC conjunction with antibiotic therapy. The mutant SPE-A is
CC especially useful for treating T cell lymphomas, and ovarian and
CC uterine cancer. It is thought that mutant SPE-A can be selectively
CC toxic to T cell lymphoma cells.
CC N.B. Sequence not given in the specification, but constructed
CC using the wild type SPE-A sequence given on pages 77-79.
XX
SQ Sequence 251 AA;
```

```
Query Match 22.3%; Score 271.5; DB 18; Length 251;
Best Local Similarity 32.5%; Pred. No. 3.4e-19;
Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;
Qy 4 SEETNEKDLRKSELOGTAL-GNLKQIYY-YNSKAIT--SSEKSDAQFLTNTLLFKGFTTG 60
Db 25 SQEVAQODPDPSQLHRSLSVLQNIYFLYEGDPVTENKSVQDQLLSHLLYN---VS 81
```

QY 61 HPWYNDLLVDLGS TAATSEYEGSSVDLYGAYYGQC-AGGTPNKTACMYGGVTLHDNNRL 119
DB 82 GPNYOKLKTTELKNQEMATLFKKNVDIYGVEYHLCYLSENAERSACIYGGVTNHEGNHL 141
QY 120 TEEKKVPINLWDGQTTPIDKVKTSKEVTQELDLQARHYLHGKFGLYNSDSFGGKV 179
DB 142 EIPKKIVKVSIDGQ-SLSFD-IETNKKWTAQELDYKVRKYLTDNKQLYTNGP--SKY 197
QY 180 QRLIVFHSSEGSTVSYDLFD----AQGOYPTDLLRIYRNTTISSTLSISLYLT 232
DB 198 ETGYIKFIPKNKESEWFDFPEPTQSKY----LMIYKNETLDSNTSQIEVYLT 250

RESULT 37
AAR13209
ID AAR13209 standard; Protein; 221 AA.
XX
AC AAR13209;
XX
DT 15-OCT-1991 (first entry)
XX
DE Streptococcal pyrogenic enterotoxin A.
XX
KW SPE A; cancer treatment; pyrogen; tumouricide; scarlet fever.
XX
OS Streptococcus NY-5 strain.
XX
PN WO9110680-A.
XX
PD 25-JUL-1991.
XX
PF 17-JAN-1991; 91WO-US00342.
XX
PR 17-JAN-1990; 90US-0466577.
XX
PA (TERM/) TERMAN D S.
XX
PI Terman DS;
XX
DR WPI; 1991-237984/32.
XX
PT Treating cancer with enterotoxin from Staphylococcus aureus -
PT administered by IV injection, having same tumoricidal activity
PT as Staphylococcal protein A without potential toxic reactions
XX
PS Disclosure; Fig 1: 74pp; English.
XX
CC SPE A can be used for tumoricidal treatment, esp. with a haemolysin.
CC Synthetic polypeptides having structural homology to Streptococcal
CC pyrogenic exotoxins are claimed, provided the homology includes
CC statistically significant sequence homology, alignment of Cysteine
CC residues and similar hydrophathy profiles.
XX
SQ See AAR13203-R13211.

Query Match 22.2%; Score 270.5; DB 12; Length 221;
Best Local Similarity 33.8%; Pred. No. 3.6e-19;
Matches 76; Conservative 42; Mismatches 88; Indels 19; Gaps 10;

QY 16 SELOQTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLTNTLFPKGFTTGHWPYNDLLVDLG 72
DB 7 SOLHRSSLVKLNQIYFLYEGDPVTHENVKSVQDQLSHDLIYN---VSGPNYDKLKTLEK 63
QY 73 STAATSEYEGSSVDLYGAYYGQC-AGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWI 131
DB 64 NQEMATLFKDKKNVDIYGVEYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIYVKVSI 123
QY 132 DGKQTPVPIDKVKTSKEVTQELDLQARHYLHGKFGLYNSDSFGGKVORGLIVFHSSEG 191
DB 124 DGIIQ-SLSFD-IETNKKWTAQELDYKVRKYLTDNKQLYTNGP--SKYETGYIKFIPKNK 179
QY 192 STVSYDLFD----AQGOYPTDLLRIYRNTTISSTLSISLYLT 232

DB 180 ESFWFDLPEPEPTQSKY----LMIYKNETLDSNTSQIEVYLT 220

RESULT 38
AAR45017
ID AAR45017 standard; protein; 221 AA.
XX
AC AAR45017;
XX
DT 08-JUN-1994 (first entry)
XX
DE Staphylococcal enterotoxin SPE A.
XX
KW Staphylococcal enterotoxin; SE; cancer; tumouricidal agent;
KW autoimmune disease; toxicity; Protein A; perfusion system.
XX
OS Staphylococcus aureus.
XX
PN WO9324136-A.
XX
PD 09-DEC-1993.
XX
PF 01-JUN-1993; 93WO-US05213.
XX
PR 01-JUN-1992; 92US-0891718.
XX
PA (STON/) STONE J L.
PA (TERM/) TERMAN D S.
XX
PI Stone JL, Terman DS;
XX
DR WPI; 1993-405418/50.
XX
PT Use of staphylococcal enterotoxin(s) and homologues - for
PT treating cancer in a patient or for the treatment of auto-immune
PT diseases
XX
PS Disclosure; Fig 1: 90pp; English.
XX
CC The sequences given in AAR45011-21 are Staphylococcal enterotoxins (SEs)
CC which may be used in the methods of the invention for treating cancer
CC in a patient. These SEs, and homologues of them, can be used as
CC tumoricidal agents for treating cancers and autoimmune disease.
CC They exhibit tumoricidal activity and toxicity identical to that
CC observed for the Protein A perfusion system. They may be administered
CC by i.v. injection.

Query Match 22.2%; Score 270.5; DB 14; Length 221;
Best Local Similarity 33.8%; Pred. No. 3.6e-19;
Matches 76; Conservative 42; Mismatches 88; Indels 19; Gaps 10;

QY 16 SELOQTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLTNTLFPKGFTTGHWPYNDLLVDLG 72
DB 7 SOLHRSSLVKLNQIYFLYEGDPVTHENVKSVQDQLSHDLIYN---VSGPNYDKLKTLEK 63
QY 73 STAATSEYEGSSVDLYGAYYGQC-AGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWI 131
DB 64 NQEMATLFKDKKNVDIYGVEYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIYVKVSI 123
QY 132 DGKQTPVPIDKVKTSKEVTQELDLQARHYLHGKFGLYNSDSFGGKVORGLIVFHSSEG 191
DB 124 DGIIQ-SLSFD-IETNKKWTAQELDYKVRKYLTDNKQLYTNGP--SKYETGYIKFIPKNK 179
QY 192 STVSYDLFD----AQGOYPTDLLRIYRNTTISSTLSISLYLT 232

DB 180 ESFWFDLPEPEPTQSKY----LMIYKNETLDSNTSQIEVYLT 220

RESULT 39
ABB76240

ABB76240 standard; Protein; 221 AA.

ABB76240;

09-AUG-2002 (first entry)

Staphylococcus pyogenes exotoxin A.

Exotoxin A; SPE A; superantigen; antigen; tumour; cancer; antitumour; therapy.

Streptococcus pyogenes.

US2002051765-A1.

02-MAY-2002.

19-DEC-2000: 2000US-0741503

31-TAN-1994. 94US-0189424

19-JUN-1995; 95US-0491746.
03-OCT-1989; 89US-0416530

17-JAN-1990; 90US-0466577.
17-JAN-1991; 91US-0466577.

01-JUN-1992; 92US-0891718.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

6

comprises

5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843

The present sequence is the prote

stretches of sequence, between staphylococcal enterotoxins and staphylococcal pyrogenic exotoxins and staphylococcal exfoliative toxins (see AB76234-44). In the present invention, synthetic polypeptides useful in tumour therapy and in blocking or destroying autoreactive T and B lymphocyte populations are characterised by substantial structural homology to staphylococcal enterotoxin A and enterotoxin B, and to streptococcal pyrogenic exotoxins, with statistically significant sequence homology and similarity (z value of Lipman and Pearson algorithm in Monte Carlo analysis exceeding 6) to include alignment of cysteine residues and similar hydrophathy profiles. These superantigens are used to treat solid tumours, including their metastases, without radiation, surgery or standard chemotherapeutic agents. A claimed method of human cancer treatment involves contacting haematopoietic cells from a patient with one or more superantigens *ex vivo* to generate stimulated cells, selecting a specific V beta subset of cells, and reintroducing these cells into the patient to induce an *in vivo* therapeutic, tumoricidal reaction.

Sequence 221 AA:

every Match	22.2%;	Score 270.5;	DB 23;	Length 221;
1st Local Similarity	33.8%;	Pred. No. 3.6e-19;		
atches 76;	Conservative 42;	Mismatches 88;	Indels 19;	Gaps 10;

16 SELOQTAL-GNLKQIYY-YNSKAIT-SSEKSADOFLTNTLLFKGFFTGHPWYNDLLVDLG 72

7 SQLHRSSLVKNLQNIYFLYEGDPVTHENVKSDQLLSHDLIYN---VSGPNYDKLKTCLK 63

73 STAATSEYEGSSVDLYGAYGYQC-AGGTPNKTACMYGGVTLHDNNRLTEEEKVPINLWI 131

64 NQEMATLFKDKNVDIYGVEYYHLCYLCENAERSACIYGGVTNHEGNHLEIPKKIVVKVSI 123

Query Match 22.2%; Score 270.5; DB 18; Length 251;
Best Local Similarity 32.5%; Pred. No. 4.3e-19;
Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SEEINEKDLRKSELOCTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLNTLLFKGFFTG 60
Db 25 SOEVFAQQDDPSQLHRSSLVNQLNIYFLYEGDPVTHENVKSVDDLSSHIIYN--VS 81
QY 61 HPWYNDDLVLGTAATSEYEGSSVDLYGAYYQ-CAGTGNKTKACMYGGVTLHDNNRL 119
Db 82 GPNYDKLTELKNOEMATLFDKKNVDIYGVEYHLSYLCENAEARSACIYGGVTNHEGNNH 141
QY 120 TEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKV 179
Db 142 EIPKKIIVKVSIDGQ-SLSFD-IETNKKMVTQAELDYKVRKYLTDNKQLYTNGP--SKY 197
QY 180 ORGLIVFHSSEGSVSVDLFD---AOGQYPTDLLRIYRDNTTISSTLSISLYLT 232
Db 198 ETGYIKFIPKNKESFWDFPEFTQSKY----LMIYKDNETLDSNTSQIEVYLTT 250

RESULT 41
AAW12151
ID AAW12151 standard; Protein; 251 AA.
AC AAW12151;
DT 04-NOV-1997 (first entry)
XX Streptococcus pyogenes Streptococcal toxin A mutant Cys87Ser.
DE Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
KW vaccine; protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
KW fever; hypotension; group A streptococcal infection; myositis;
KW fascitis; liver damage; T cell; lymphoma; ovarian; uterine.
XX Streptococcus pyogenes.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT Peptide /label= sig_peptide
FT Peptide 31..251
FT Misc-difference 117 /label= mat_peptide
FT /note= "wild type Cys replaced by Ser"
XX
PN WO9640930-Al.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US10252.
XX
XX 07-JUN-1995; 95US-0480261.
XX
XX (MINU) UNIV MINNESOTA.
XX
XX Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
XX WPI; 1997-099936/09.
XX
XX Mutant SPE-A toxin with at least one amino acid change is
PT substantially non-lethal - used in vaccine composition for
PT treatment of cancer and streptococcal toxic shock syndrome etc.
XX
XX Example 4; Page -: 102pp; English.
XX
XX The present sequence is a non-lethal Streptococcus pyogenes
CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce
CC vaccines to protect animals against wild type SPE-A and to treat
CC cancer and streptococcal toxic shock syndrome (STSS). The mutant

CC SPE-A causes neutralising antibodies (Ab) to be produced, which
CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,
CC group A streptococcal infection, myositis, fascitis and liver
CC damage. The neutralising Ab is preferably administered in
CC conjunction with antibiotic therapy. The mutant SPE-A is
CC especially useful for treating T cell lymphomas, and ovarian and
CC uterine cancer. It is thought that mutant SPE-A can be selectively
CC toxic to T cell lymphoma cells.
CC N.B. Sequence not given in the specification, but constructed
CC using the wild type SPE-A sequence given on pages 77-79.
XX
SQ Sequence 251 AA;

Query Match 22.2%; Score 270.5; DB 18; Length 251;
Best Local Similarity 32.6%; Pred. No. 4.3e-19;
Matches 78; Conservative 44; Mismatches 94; Indels 23; Gaps 11;

QY 4 SEEINEKDLRKSELOCTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLNTLLFKGFFTG 60
Db 25 SOEVFAQQDDPSQLHRSSLVNQLNIYFLYEGDPVTHENVKSVDDLSSHIIYN--VS 81
QY 61 HPWYNDDLVLGTAATSEYEGSSVDLYGAYY---GYOCAGGTPNKTACMYGGVTLHDNN 117
Db 82 GPNYDKLTELKNOEMATLFDKKNVDIYGVEYHLSYLCENAEARSACIYGGVTNHEGNN 139
QY 118 RLTEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGG 177
Db 140 HLEIPKKIIVKVSIDGQ-SLSFD-IETNKKMVTQAELDYKVRKYLTDNKQLYTNGP--S 195
QY 178 KVQGLIVFHSSEGSVSVDLFD---AOGQYPTDLLRIYRDNTTISSTLSISLYLT 232
Db 196 KYETGYIKFIPKNKESFWDFPEFTQSKY----LMIYKDNETLDSNTSQIEVYLTT 250

RESULT 42
AAW12154
ID AAW12154 standard; Protein; 251 AA.
XX
AC AAW12154;
XX
DT 04-NOV-1997 (first entry)
XX
DE Streptococcus pyogenes Streptococcal toxin A mutant Ser195Ala.
XX
KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
KW vaccine; protection; treatment; cancer; neutralising antibody;
KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
KW fever; hypotension; group A streptococcal infection; myositis;
KW fascitis; liver damage; T cell; lymphoma; ovarian; uterine.
XX Streptococcus pyogenes.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 1..30
FT Peptide /label= sig_peptide
FT Peptide 31..251
FT Misc-difference 225 /label= mat_peptide
FT /note= "wild type Ser replaced by Ala"
XX
PN WO9640930-Al.
XX
XX 19-DEC-1996.
XX
XX 07-JUN-1996; 96WO-US10252.
XX
XX 07-JUN-1995; 95US-0480261.
XX
XX (MINU) UNIV MINNESOTA.
XX
XX Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
XX

DR WPI; 1997-099936/09.

XX Mutant SPE-A toxin with at least one amino acid change is

PT substantially non-lethal - used in vaccine composition for

PT treatment of cancer and streptococcal toxic shock syndrome etc.

XX

XX Example 4; Page -: 102pp; English.

XX

XX The present sequence is a non-lethal Streptococcus pyogenes

CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce

CC vaccines to protect animals against wild type SPE-A and to treat

CC cancer and streptococcal toxic shock syndrome (STSS). The mutant

CC SPE-A causes neutralising antibodies (Ab) to be produced, which

CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,

CC group A streptococcal infection, myositis, fasciitis and liver

CC damage. The neutralising Ab is preferably administered in

CC conjunction with antibiotic therapy. The mutant SPE-A is

CC especially useful for treating T cell lymphomas, and ovarian and

CC uterine cancer. It is thought that mutant SPE-A can be selectively

CC toxic to T cell lymphoma cells.

CC N.B. Sequence not given in the specification, but constructed

CC using the wild type SPE-A sequence given on pages 77-79.

XX

XX Sequence 251 AA;

CC

Query Match 22.0%; Score 268.5; DB 18; Length 251;

Best Local Similarity 32.5%; Pred. No. 6.9e-19;

Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SBEINEKLRKSELGQTAL-GNLKOIYY-YNSKAIT-SSEKSADQFLTNTLLKGFFTG 60

Db 25 SGEVFAQQDPPSQLHRSLSLVNQLNIYFLYEGDPVTHENVKSVQDLSHLLIYN---VS 81

QY 61 HPWYNDLLVDLGSTAASTSEYEGSSVDLYGAYGYOC-AGGTPNKACMYGGVTLHDNNRL 119

Db 82 GPNYDKLTKELKNQEMATLFDKNDVIYGVVEYHLCYLCEAERSACIYGGVTNHEGNHL 141

QY 120 TEEKKVPINLWDGKQTTVPIDKVKTSKKEVTQVELDQARHYLHGKFLGYNDSDFGSKV 179

Db 142 EIPKKIVKVSIDGIQ-SLSFD-IETNKKMVTQALDYKVKYLDNDKQLYTNGP--SKY 197

QY 180 QRLIVFHSSEGSTVSYDLFD----AQQYPTLLRIYRDNNTTISSTLSISLYLT 232

Db 198 ETGYIKFIPKNKESFWDFPEPEFTQAKY----LMIYKDNELTSDNTSQIEVYLTT 250

RESULT 43

AAW12148

ID AAW12148 standard; Protein; 251 AA.

XX

XX AAW12148;

XX

XX 04-NOV-1997 (first entry)

XX

XX Streptococcus pyogenes Streptococcal toxin A mutant N20D/K157E.

DE

XX Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;

KW vaccine; protection; treatment; cancer; neutralising antibody;

KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;

KW fever; hypotension; group A streptococcal infection; myositis;

KW fasciitis; liver damage; T cell; lymphoma; ovarian; uterine.

XX

XX Streptococcus pyogenes.

OS

OS Synthetic.

XX

XX Key Location/Qualifiers

FT Peptide 1..30

FT /label= sig_peptide

FT Peptide 31..251

FT /label= mat_peptide

FT Misc-difference 50

FT /note= "wild type Asn replaced by Asp"

FT Misc-difference 187

FT /note= "wild type Lys replaced by Glu"

XX

PN WO9640930-A1.

XX

PD 19-DEC-1996.

XX

PF 07-JUN-1996; 96WO-US10252.

XX

XX 07-JUN-1995; 95US-0480261.

XX

PA (MINU) UNIV MINNESOTA.

XX

PI Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;

XX

DR WPI; 1997-099936/09.

XX

XX Mutant SPE-A toxin with at least one amino acid change is

PT substantially non-lethal - used in vaccine composition for

PT treatment of cancer and streptococcal toxic shock syndrome etc.

XX

PS Claim 5; Page -: 102pp; English.

XX

XX The present sequence is a non-lethal Streptococcus pyogenes

CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce

CC vaccines to protect animals against wild type SPE-A and to treat

CC cancer and streptococcal toxic shock syndrome (STSS). The mutant

CC SPE-A causes neutralising antibodies (Ab) to be produced, which

CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,

CC group A streptococcal infection, myositis, fasciitis and liver

CC damage. The neutralising Ab is preferably administered in

CC conjunction with antibiotic therapy. The mutant SPE-A is

CC especially useful for treating T cell lymphomas, and ovarian and

CC uterine cancer. It is thought that mutant SPE-A can be selectively

CC toxic to T cell lymphoma cells.

CC N.B. Sequence not given in the specification, but constructed

CC using the wild type SPE-A sequence given on pages 77-79.

XX

XX Sequence 251 AA;

CC

Query Match 22.0%; Score 268.5; DB 18; Length 251;

Best Local Similarity 32.5%; Pred. No. 6.9e-19;

Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SBEINEKLRKSELGQTAL-GNLKOIYY-YNSKAIT-SSEKSADQFLTNTLLKGFFTG 60

Db 25 SGEVFAQQDPPSQLHRSLSLVNQLNIYFLYEGDPVTHENVKSVQDLSHLLIYN---VS 81

QY 61 HPWYNDLLVDLGSTAASTSEYEGSSVDLYGAYGYOC-AGGTPNKACMYGGVTLHDNNRL 119

Db 82 GPNYDKLTKELKNQEMATLFDKNDVIYGVVEYHLCYLCEAERSACIYGGVTNHEGNHL 141

QY 120 TEEKKVPINLWDGKQTTVPIDKVKTSKKEVTQVELDQARHYLHGKFLGYNDSDFGSKV 179

Db 142 EIPKKIVKVSIDGIQ-SLSFD-IETNKKMVTQALDYKVKYLDNDKQLYTNGP--SKY 197

QY 180 QRLIVFHSSEGSTVSYDLFD----AQQYPTLLRIYRDNNTTISSTLSISLYLT 232

Db 198 ETGYIKFIPKNKESFWDFPEPEFTQSKY----LMIYKDNELTSDNTSQIEVYLTT 250

RESULT 44

AAW12147

ID AAW12147 standard; Protein; 251 AA.

XX

XX AAW12147;

XX

XX 04-NOV-1997 (first entry)

XX

XX Streptococcus pyogenes Streptococcal toxin A mutant Asn20Asp.

XX

XX Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;

KW vaccine; protection; treatment; cancer; neutralising antibody;

KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;

KW fever; hypotension; group A streptococcal infection; myositis;
 KW fasciitis; liver damage; T cell; lymphoma; ovarian; uterine.
 XX Streptococcus pyogenes.
 OS Synthetic.

XX Key Location/Qualifiers
 FH Peptide 1..30
 FT /label= sig_peptide
 FT Peptide 31..251
 FT /label= mat_peptide
 FT Misc-difference 50
 FT /note= "wild type Asn replaced by Asp"

XX WO9640930-A1.

XX 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US10252.

XX PR 07-JUN-1995; 95US-0480261.

XX PA (MINU) UNIV MINNESOTA.

XX PI Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;

XX WPI; 1997-099936/09.

XX Mutant SPE-A toxin with at least one amino acid change is
 PT substantially non-lethal - used in vaccine composition for
 PT treatment of cancer and streptococcal toxic shock syndrome etc.

XX Claim 5; Page -: 102pp; English.

XX The present sequence is a non-lethal Streptococcus pyogenes
 CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce
 CC vaccines to protect animals against wild type SPE-A and to treat
 CC cancer and streptococcal toxic shock syndrome (STSS). The mutant
 CC SPE-A causes neutralising antibodies (Ab) to be produced, which
 CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,
 CC group A streptococcal infection, myositis, fasciitis and liver
 CC damage. The neutralising Ab is preferably administered in
 CC conjunction with antibiotic therapy. The mutant SPE-A is
 CC especially useful for treating T cell lymphomas, and ovarian and
 CC uterine cancer. It is thought that mutant SPE-A can be selectively
 CC toxic to T cell lymphoma cells.
 CC N.B. Sequence not given in the specification, but constructed
 CC using the wild type SPE-A sequence given on pages 77-79.

XX Sequence 251 AA;

Query Match 22.0%; Score 268.5; DB 18; Length 251;
 Best Local Similarity 32.5%; Pred. No. 6.9e-19;
 Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SEINEKDLRKXSELOGTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLNTLLFGFFTG 60

Db 25 SOEVAQQDPDSQLHRSSLVKNLQDIPLYEGDPVTHENVKSVQLLSHLLYN--VS 81

QY 61 HPWYNDLLVDLGLSTAAATSEYEGSSVDLYGAYYGYOC-AGGTPNKTCMYGGVTLHDNNRL 119

Db 82 GPNYDKLTKELKNQEMATLFDKNDVIYGVYHYLCYLCENASACIYGGVTNHEGNHL 141

QY 120 TEEKKVPINLWDIGKOTVPIDKVTSKKEVTQVLELDLQARHLHGKFLGNSDFGGKV 179

Db 142 ETPKILVAVKVSIDGQI-Q-SUSFD-IETNKKMVAQELDYKRVKLYTDNKKQLYTNGP--SKY 197

QY 180 QRGLIVFHSSEGSTVSYDLFD----AQQGYPTLLRIYRDNTTISSTLSISLYLT 232

Db 198 ETGYIKFIPKNKESWFDFEPPEFTQSKY----LMYKDNFTLDSNTSQIEVYLT 250

RESULT 45

AAW12146

XX ID AAW12146 standard; Protein; 251 AA.

XX AC AAW12146;

XX DF 04-NOV-1997 (first entry)

XX Streptococcus pyogenes Streptococcal toxin A mutant Lys157Glu.

XX Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
 KW vaccine; protection; treatment; cancer; neutralising antibody;
 KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
 KW fever; hypotension; group A streptococcal infection; myositis;
 KW fasciitis; liver damage; T cell; lymphoma; ovarian; uterine.

XX Streptococcus pyogenes.

XX Synthetic.

XX Key Location/Qualifiers

FT Peptide 1..30

FT /label= sig_peptide

FT Peptide 31..251

FT /label= mat_peptide

FT Misc-difference 187

FT /note= "wild type Lys replaced by Glu"

XX WO9640930-A1.

XX 19-DEC-1996.

XX PF 07-JUN-1996; 96WO-US10252.

XX PR 07-JUN-1995; 95US-0480261.

XX (MINU) UNIV MINNESOTA.

XX Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;

XX WPI; 1997-099936/09.

XX Mutant SPE-A toxin with at least one amino acid change is
 PT substantially non-lethal - used in vaccine composition for
 PT treatment of cancer and streptococcal toxic shock syndrome etc.

XX Claim 5; Page -: 102pp; English.

XX The present sequence is a non-lethal Streptococcus pyogenes
 CC Streptococcal toxin A (SPE-A) mutant, which can be used to produce
 CC vaccines to protect animals against wild type SPE-A and to treat
 CC cancer and streptococcal toxic shock syndrome (STSS). The mutant
 CC SPE-A causes neutralising antibodies (Ab) to be produced, which
 CC may be used to ameliorate STSS symptoms, e.g. fever, hypotension,
 CC group A streptococcal infection, myositis, fasciitis and liver
 CC damage. The neutralising Ab is preferably administered in
 CC conjunction with antibiotic therapy. The mutant SPE-A is
 CC especially useful for treating T cell lymphomas, and ovarian and
 CC uterine cancer. It is thought that mutant SPE-A can be selectively
 CC toxic to T cell lymphoma cells.
 CC N.B. Sequence not given in the specification, but constructed
 CC using the wild type SPE-A sequence given on pages 77-79.

XX Sequence 251 AA;

Query Match 22.0%; Score 268.5; DB 18; Length 251;
 Best Local Similarity 32.5%; Pred. No. 6.9e-19;
 Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SEINEKDLRKXSELOGTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLNTLLFGFFTG 60

Db 25 SOEVAQQDPDSQLHRSSLVKNLQDIPLYEGDPVTHENVKSVQLLSHLLYN--VS 81

QY 61 HPWYNDLLVDLGLSTAAATSEYEGSSVDLYGAYYGYOC-AGGTPNKTCMYGGVTLHDNNRL 119

Db 82 GPNYDKLTELKQEMATLFDKDNVDIYGVYHLYCENAEARSACIYGGVTNHEGHL 141
 QY 120 TEKKVPINLWIDGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFLGNSDSFGGKV 179
 Db 142 EIPKKIVVKVSDIGIQ-SLSFD-IETNKKMVTQAQELDKVRKYLTDNEQLYNGP--SKY 197
 QY 180 QRLIVFHSSEGSTVSYDLFD----AQOYPTLLRIYRDNTTISSTLSISLYLYT 232
 Db 198 ETGYIKFIPKNKSEFWDFPEFTQSKY----LMIYKDNELTDSNTSQIEVYLT 250

RESULT 46
 AAW12097
 ID AAW12097 standard; Protein: 251 AA.
 XX
 AC AAW12097;
 XX
 DT 04-NOV-1997 (first entry)
 XX
 DE Streptococcus pyogenes Streptococcal toxin A.
 XX
 KW Streptococcal; toxin A; SPE-A; non-lethal; mutant; production;
 KW vaccine; protection; treatment; cancer; neutralising antibody;
 KW streptococcal toxic shock syndrome; STSS; symptom; amelioration;
 KW fever; hypotension; group A streptococcal infection; myositis;
 KW fasciitis; liver damage; T cell; lymphoma; ovarian; uterine.
 XX
 OS Streptococcus pyogenes.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT /label= sig_peptide
 FT Peptide 31..251
 FT /label= mat_peptide
 XX
 PN W09640930-A1.
 XX
 PD 19-DEC-1996.
 XX
 PF 07-JUN-1996; 96WO-US10252.
 XX
 PR 07-JUN-1995; 95US-0480261.
 XX
 PA (MINU) UNIV MINNESOTA.
 XX
 PI Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
 XX
 DR WPI: 1997-099936/09.
 DR N-PSDB; AAW12097.
 XX
 PT Mutant SPE-A toxin with at least one amino acid change is
 PT substantially non-lethal - used in vaccine composition for
 PT treatment of cancer and streptococcal toxic shock syndrome etc.
 XX
 PS Disclosure: Pages 77-79; 102pp: English.
 XX
 CC The present sequence is Streptococcus pyogenes Streptococcal
 CC toxin A (SPE-A), from which a non-lethal mutant SPE-A, comprising
 CC at least 1 amino acid change, can be derived. The mutant SPE-A can
 CC be used to produce vaccines to protect animals against wild type
 CC SPE-A and to treat cancer and streptococcal toxic shock syndrome
 CC (STSS). The mutant SPE-A causes neutralising antibodies (Ab) to be
 CC produced, which may be used to ameliorate STSS symptoms, e.g.
 CC fever, hypotension, group A streptococcal infection, myositis,
 CC fasciitis and liver damage. The neutralising Ab is preferably
 CC administered in conjunction with antibiotic therapy. The mutant
 CC SPE-A is especially useful for treating T cell lymphomas, and
 CC ovarian and uterine cancer. It is thought that mutant SPE-A can be
 CC selectively toxic to T cell lymphoma cells.
 XX
 SQ Sequence 251 AA;

Query Match 22.0%; Score 268.5; DB 18; Length 251;

Best Local Similarity 32.5%; Pred. No. 6.9e-19;
 Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;
 QY 4 SBEINEKDLRKKSELOCTAL-GNLKQIYY-YNSKAIT-SSEKSADOFLLNTLLFGFFTG 60
 Db 25 SOEFAAQDPDPQSILHRSLSVKNLQNIYFLYEGDPVTHENKSVQDLSSHLLIYN---VS 81
 QY 61 HPWYNDLVLIDLGSTAATSEYEGSSVDLYGAYGYOC-AGGTPNKTKACMYGGVTLHDNNRL 119
 Db 82 GPNYDKLTELKQEMATLFDKDNVDIYGVYHLYCENAEARSACIYGGVTNHEGHL 141
 QY 120 TEKKVPINLWIDGKQTTVPIDKVKTSKEVTQVQELDLQARHYLHGKFLGNSDSFGGKV 179
 Db 142 EIPKKIVVKVSDIGIQ-SLSFD-IETNKKMVTQAQELDKVRKYLTDNQLYNGP--SKY 197
 QY 180 QRLIVFHSSEGSTVSYDLFD----AQOYPTLLRIYRDNTTISSTLSISLYLYT 232
 Db 198 ETGYIKFIPKNKSEFWDFPEFTQSKY----LMIYKDNELTDSNTSQIEVYLT 250

RESULT 47

AAW59780
 ID AAW59780 standard; Protein: 251 AA.
 XX
 AC AAW59780;
 XX
 DT 12-OCT-1998 (first entry)
 XX
 DE Amino acid sequence of Streptococcus pyogenes exotoxin A.
 XX
 KW SPE-A toxin; nonlethal; mutant; Streptococcus pyogenes exotoxin A;
 KW wild type; nontoxic; antibody; vaccine; immunity; ovarian cancer;
 KW streptococcal toxic shock syndrome; STSS; T cell lymphoma;
 KW uterine cancer.
 XX
 OS Streptococcus pyogenes.
 XX
 PN W09824911-A2.
 XX
 PD 11-JUN-1998.
 XX
 PF 05-DEC-1997; 97WO-US222228.
 XX
 PR 06-DEC-1996; 96US-0032930.
 XX
 PA (MINU) UNIV MINNESOTA.
 XX
 PI Ohlendorf D, Roggiani M, Schlievert PM, Stoehr J;
 XX
 DR WPI: 1998-333330/29.
 DR N-PSDB; AAW41593.
 XX
 PT New mutant Streptococcal SPE-A toxins - useful for, e.g. prevention
 PT or treatment of streptococcal infection or toxic shock syndrome
 XX
 PS Disclosure: Fig 3; 95pp: English.
 XX
 CC This is the amino acid sequence of the Streptococcus pyogenes exotoxin A
 CC (SPE-A toxin). The novel mutant Streptococcal SPE-A toxin has at least
 CC 1 aa change and is nonlethal compared with a protein to wild type SPE-A
 CC toxin. The mutant SPE-A toxins are nontoxic and can produce antibodies
 CC that neutralise wild type SPE-A toxin activity. The toxins can be used
 CC in vaccines and therapeutics to generate a protective immune response
 CC against streptococcal infection. They can be used to protect against the
 CC development of streptococcal toxic shock syndrome (STSS). In addition,
 CC the toxins can be used for treating animals with symptoms of
 CC streptococcal infection or STSS and in methods for stimulating T cell
 CC proliferation and in the treatment of cancer. In particular they can be
 CC used for treating T cell lymphomas, and ovarian and uterine cancer.
 XX
 SQ Sequence 251 AA;

Query Match 22.0%; Score 268.5; DB 19; Length 251;

Best Local Similarity 32.5%; Pred. No. 6.9e-19;
Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SEINEKDLRKSELOGTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLNTLLFKGFFTC 60
Db 25 SQVFQAQDDPSQLHRSLVKNLQNIYFLYEGDPVTHENVKSDQLLSHLLYN---VS 81
QY 61 HPYNDLLVDLGSTAATSEYEGSSVDLYGAYGYQC-AGGTPNKTCMYGGVTLHDNNRL 119
Db 82 GPNYDKLTQELKNQEMATLFKDNVDIYVEYHLCYLCEAERSACIYGGVTNHEGNHL 141
QY 120 TEKKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDSFGKV 179
Db 142 EIPKKIWKVVSIDGIO-SLSFD-IETNKKMVTQAQELDYKVRKYLTDNKQLYTNGP--SKY 197
QY 180 ORGLIVFHSSEGSTVSYDLFD---AQGYPDVTLRLIYRDNTTISSTLSLSLYLT 232
Db 198 ETGYIKFIPKNKESFDFPEPEFTOSKY----LMIYKDNETLDSNTSQIEVILTT 250

RESULT 48
AAW64647
ID AAW64647 standard; peptide; 239 AA.
XX
AC AAW64647;
XX
DT 23-OCT-1998 (first entry)
XX
DE Synthetic SEB protein fragment.
XX
KW Enterotoxin B; SEB; pyrogenic exotoxin; PET; protective immunity;
KW toxic shock; toxin-mediated activation; T-cell; antagonist; inhibitor;
KW therapeutic; vaccine; food poisoning.
XX
XX Synthetic.
OS Staphylococcus aureus.
XX
XX WO9829444-A1.
XX
XX 09-JUL-1998.
XX
XX 30-DEC-1997; 97WO-TL00438.
XX
XX 30-DEC-1996; 96IL-0119938.
XX
XX (YISS) YISSUM RES & DEV CO.
XX
XX Arad G, Kaempfer R;
XX
XX WPI; 1998-388042/33.
XX
XX New peptide(s) derived from pyrogenic exotoxin - useful for, e.g.
XX antagonising toxin-mediated activation of T cells and prevention or
XX treatment of toxic shock caused by exotoxin(s)
XX
XX Example 8; Page 41; 68pp; English.
XX
XX AAW64636-W64657 are peptides homologous to the amino acid sequence of a
XX fragment of a pyrogenic exotoxin (PET), and derivatives of the peptide
XX capable of eliciting protective immunity against toxic shock induced by
XX PET or by a mixture of PETs. Such peptides are also capable of
XX antagonising toxin-mediated activation of T-cells, inhibiting expression
XX of pyrogenic toxin (PT)-induced mRNA encoded by IL-2, IFN-gamma or
XX TNF-beta genes. The peptides may be used to prepare therapeutics or
XX vaccines for the treatment of prophylaxis of toxin-mediated activation
XX of T cells and eliciting protective immunity against toxic shock induced
XX by PETs. They can also be used for the treatment of harmful effects
XX (especially food poisoning) and toxic shock caused by PET. Antiserum to
XX the peptides can also be used for alleviating toxic shock induced by
XX PET.

Sequence 239 AA;

Query Match 22.0%; Score 267.5; DB 19; Length 239;
Best Local Similarity 31.7%; Pred. No. 8.1e-19;
Matches 77; Conservative 43; Mismatches 104; Indels 19; Gaps 8;

QY 2 EKSEETNEKDLRKSELOGTALGNLKOIYYNSKAITTSSEKSADQFLNTLLFKGFFTC 61
Db 1 ESQDPDKPDLKHKSSKFTG-LMENMK-VLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKL 58
QY 62 PWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTCMYGGV 111
Db 59 GNYDNVRFKKNKDLADKYDKYDVDFGANYYYOCYFSKKTNDINSHETDKRKTCTMYGV 118
QY 112 TLHDNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYN 171
Db 119 TEHNGNQLDKYRSITVRVFDGK-NLLSFD-VQTNKKKVTQAQELDYLRHYLVKNKKLYE 176
QY 172 SDSFGKVGRLIVFHSSEGSTVSYDLFDAQGOYPD--TLLRIYRDNTTISSTLSLSLY 229
Db 177 FNN--SPYETGYIKFIENENS-FWYDMPAPGDKFDQSKYLMYNDKNKMYDSKDVKIEVY 233

QY 230 LYT 232
Db 234 LYT 236

RESULT 49
AAB67341
ID AAB67341 standard; peptide; 239 AA.
XX
AC AAB67341;
XX
DT 23-APR-2001 (first entry)
XX
DE Staphylococcus aureus enterotoxin B protein.
XX
XX Tumour; cancer; immune; enterotoxin.
XX
XX Staphylococcus aureus.
XX
XX US6180097-B1.
XX
XX 30-JAN-2001.
XX
XX 30-OCT-1998; 98US-0183437.
XX
XX 31-JAN-1994; 94US-0189424.
XX 19-JUN-1995; 95US-0491746.
XX 03-OCT-1989; 89US-0416530.
XX 17-JAN-1990; 90US-0466577.
XX 17-JAN-1991; 91WO-US00342.
XX 01-JUN-1992; 92US-0891718.
XX 02-MAR-1993; 93US-0025144.
XX (TERM/) Terman D S.
XX
XX Terman DS;
XX
XX WPI; 2001-158657/16.
XX
XX Tumor cell capable of stimulating antitumor immune reactivity in vitro
XX or in vivo comprises exogenous nucleic acids encoding a superantigen
XX and a costimulatory molecule -
XX
XX Disclosure; Fig 2; 16pp; English.
XX
XX The present invention relates to a tumour cell capable of stimulating
XX antitumor immune reactivity in vitro or in vivo contains and
XX expresses an exogenous nucleic acid molecule encoding a superantigen
XX or its active fragment and an exogenous nucleic acid molecule
XX encoding a costimulatory molecule that activates T cells in
XX conjunction with an antigenic stimulus. The invention may be used
XX for cancer therapy by stimulating an anticancer immune response
XX in vivo or ex vivo.

```
XX Sequence 239 AA;
SQ Query Match 22.0%; Score 267.5; DB 22; Length 239;
Best Local Similarity 31.7%; Pred. No. 8.1e-19;
Matches 77; Conservative 43; Mismatches 104; Indels 19; Gaps 8;

QY 2 EKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSEKSAQDPLTWTLLFKGFFTGH 61
DQ 1 ESQDPKPDDELHKSSKFTG-LMENMK-VLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKL 58

QY 62 PWYNDLLVGLGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTCMYGV 111
DQ 59 GNYDNRVVEFKKDLADKYDKYVDFGANYYYQCYFSKKTNDINSHQTDKRTKTCMYGV 118

QY 112 TLHDNNRLTEKKVPIINLWIDGKQITVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYN 171
DQ 119 TEHNGNQLDKYRSITVRVFEDEGK-NLLSPD-VQTNKKVTAQELDYLTRHYLVKNKKLYE 176

QY 172 SDSFGKVGQGLIVFHSSEGSTVSVDLFAQOQYDP--TLLRIYRDNNTTISSTLSISLY 229
DQ 177 FNN--SPYETGYIKFIENENS-FWYDMMPPAGDKFDQSKYLMYNDNKNVDSKDYKIEVY 233

QY 230 LYT 232
DQ 234 LTT 236

RESULT 50
AAW06737
ID AAW06737 standard; Protein; 255 AA.
XX
AC AAW06737;
DT 08-MAR-1997 (first entry)
XX
DE Staphylococcus enterotoxin B.
XX
KW Enterotoxin B; superantigen; antigen; cytokine; chemokine; T cell;
KW lymphocyte; monocyte; natural killer cell; gene therapy; cancer;
KW vaccine; adjuvant.
XX
OS Staphylococcus sp.
XX
FH Key Location/Qualifiers
FT Peptide 1..15
FT /label= Sig_peptide
XX
PN W09G36366-A1.
XX
PD 21-NOV-1996.
XX
PF 20-MAY-1996; 96WO-US07432.
XX
PR 29-DEC-1995; 95US-0580806.
PR 18-MAY-1995; 95US-0446918.
XX
PA (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY MED.
XX
PI Dow SW, Elmslie RE, Potter TA;
XX
WPI; 1997-011857/01.
DR N-PSDB; AAT45698.
XX
PT Recombinant molecule encoding superantigen and opt. cytokine or
PT chemokine - controls activity of effector cells (T cells, monocytes,
PT natural killer cells), used for gene therapy of cancer
XX
PS Example 1; Page 96-97; 131pp; English.
XX
CC A cDNA clone (AAT45698) codes for staphylococcal enterotoxin B
CC (AAW06737) superantigen. Nucleic acids encoding superantigens (see
CC also AAW06738-39), esp. truncated forms of the superantigen lacking
```

```
CC the leader peptide, can be utilised in the gene therapy of cancer,
CC infectious diseases and immunological disorders. The nucleic acid,
CC optionally in combination with cytokine or chemokine nucleic acids,
CC is delivered to an animal using e.g. liposomes. It acts by
CC controlling the activity of effector cells, such as T-cells,
CC macrophages, monocytes and/or natural killer cells. Localised
CC prodn. of an effective but non-toxic amount of encoded proteins
CC allows safe treatment of the animal.
XX
SQ Sequence 255 AA;

Query Match 22.0%; Score 267.5; DB 18; Length 255;
Best Local Similarity 31.7%; Pred. No. 8.9e-19;
Matches 77; Conservative 43; Mismatches 104; Indels 19; Gaps 8;

QY 2 EKSEINEKDLRKKSELOGTALGNLKOIYYNSKAITSEKSAQDPLTWTLLFKGFFTGH 61
DQ 17 ESQDPKPDDELHKSSKFTG-LMENMK-VLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKL 74

QY 62 PWYNDLLVGLGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTCMYGV 111
DQ 75 GNYDNRVVEFKKDLADKYDKYVDFGANYYYQCYFSKKTNDINSHQTDKRTKTCMYGV 134

QY 112 TLHDNNRLTEKKVPIINLWIDGKQITVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYN 171
DQ 135 TEHNGNQLDKYRSITVRVFEDEGK-NLLSPD-VQTNKKVTAQELDYLTRHYLVKNKKLYE 192

QY 172 SDSFGKVGQGLIVFHSSEGSTVSVDLFAQOQYDP--TLLRIYRDNNTTISSTLSISLY 229
DQ 193 FNN--SPYETGYIKFIENENS-FWYDMMPPAGDKFDQSKYLMYNDNKNVDSKDYKIEVY 249

QY 230 LYT 232
DQ 250 LTT 252

Search completed: January 8, 2003, 11:56:54
Job time : 21.4923 secs
```

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: January 8, 2003, 11:55:37 ; Search time 6.95138 Seconds
(without alignments)
986.213 Million cell updates/sec

Title: US-09-900-766-2

Perfect score: 1218

Sequence: 1 SEKSEEINEXDLRKKSELQ.....RDNTTISISLSLYLTT 233

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1107	90.9	257	3	US-08-486-099-112
2	1107	90.9	257	3	US-08-360-107A-122
3	1107	90.9	257	3	US-08-484-223B-112
4	1107	90.9	257	3	US-08-919-597-112
5	1107	90.9	257	3	US-08-475-668A-112
6	1107	90.9	257	3	US-08-485-551A-112
7	1107	90.9	257	3	US-08-471-913A-112
8	1107	90.9	257	4	US-08-485-264A-112
9	1107	85.4	226	4	US-08-474-349A-112
10	1040	85.4	226	4	US-08-896-933-24
11	1040	85.4	226	4	US-09-314-235-24
12	948	77.8	257	3	US-08-486-099-113
13	948	77.8	257	3	US-08-360-107A-123
14	948	77.8	257	3	US-08-484-223B-113
15	948	77.8	257	3	US-08-919-597-113
16	948	77.8	257	3	US-08-475-668A-113
17	948	77.8	257	3	US-08-485-551A-113
18	948	77.8	257	3	US-08-471-913A-113
19	948	77.8	257	4	US-08-485-264A-113
20	948	77.8	257	4	US-08-474-349A-113
21	944	77.5	233	1	US-08-446-918A-4
22	944	77.5	233	2	US-08-580-806-4
23	931.5	76.5	232	4	US-08-896-933-23
24	931.5	76.5	232	4	US-09-314-235-23
25	925	75.9	257	4	US-09-144-776B-2
26	921	75.6	233	4	US-09-144-776B-4
27	605	49.7	228	4	US-08-896-933-25

Sequence 25, Appl
Sequence 19, Appl
Sequence 17, Appl
Sequence 28, Appl
Sequence 29, Appl
Sequence 29, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 27, Appl
Sequence 16, Appl
Sequence 26, Appl
Sequence 8, Appl
Sequence 6, Appl
Sequence 10, Appl
Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-08-486-099-112
; Sequence 112, Application US/08486099
; Patent No. 6013263

; GENERAL INFORMATION:

; APPLICANT: Bolognesi, Dani P.

; APPLICANT: Matthews, Thomas J.

; APPLICANT: Wild, Carl T.

; APPLICANT: Barney, Shawn O.

; APPLICANT: Lambert, Dennis M.

; APPLICANT: Petteway, Stephen R.

; APPLICANT: Langlois, Alphonse J.

; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF

; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS

; NUMBER OF SEQUENCES: 209

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/486,099

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Coruzzi, Laura A.

; REGISTRATION NUMBER: 30,742

; REFERENCE/DOCKET NUMBER: 7872-031

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-9741/8864

; TELEFAX: 66141 PENNIE

; INFORMATION FOR SEQ ID NO: 112:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 257 amino acids

; TYPE: amino acid

; STRANDEDNESS:

; TOPOLOGY: unknown

; MOLECULE TYPE: protein

; US-08-486-099-112

Query Match 90.9%; Score 1107; DB 3; Length 257;

```

Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKSELOQTALGNLKOIYYYNKSKAITSEKSDQFLTNTLLFKGFFTG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 25 SEKSEINEKDLRKSELOQRNALSNLRIOIYYYNKAITENKESDQFLENTLLFKGFFTG 84
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 61 HPWYNDLLVDLGSAAATSEYEGSSVDLYGAYGYOCAGTGNKTCMYGGVTLDHNNRLT 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 HPWYNDLLVDLGSADATNKYKGVKVDLYGAYGYOCAGTGNKTCMYGGVTLDHNNRLT 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 121 EEKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQ 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 145 EEKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDSFGGKVQ 204
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 181 RGLVFFHSSEGSTSVSYDLFDAQGOVPTDLLRIYRDNTTISSTLSISLYLYTT 233
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 205 RGLVFFHSSEGSTSVSYDLFDAQGOVPTDLLRIYRDNTTISSTLSISLYLYTT 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
US-08-360-107A-122
: Sequence 122, Application US/08360107A
: Patent No. 6017536
: GENERAL INFORMATION:
: APPLICANT: Bolognesi, Dani P.
: APPLICANT: Matthews, Thomas J.
: APPLICANT: Wild, Carl T.
: APPLICANT: Barney, Shawn O.
: APPLICANT: Lambert, Dennis M.
: APPLICANT: Petteway, Stephen R.
: APPLICANT: Langlois, Alphonse J.
: TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
: TITLE OF INVENTION: TRANSMISSION
: NUMBER OF SEQUENCES: 149
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/360,107A
: FILING DATE: 20-DEC-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Coruzzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7872-013
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 122:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 257 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-360-107A-122

Query Match 90.9%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

```

```

QY 1 SEKSEENKDLRKRSLOQTALGNLQKQIYYNSKAITSEKSADQFLNTLLFKGFFTG 60
Db 25 SEKSEENKDLRKRSLOQTALGNLQKQIYYNSKAITSEKSADQFLNTLLFKGFFTG 84
QY 61 HPWYNLLVDLGSSTAATSEYEGSSVDLYGAYICYQACAGGTPNKTCACMYGGVTLHDNNRLT 120
Db 85 HPWYNLLVDLGSSTADATNKYGGKRVLDLYGAYICYQACAGGTPNKTCACMYGGVTLHDNNRLT 144
QY 121 EEKKVPINLWDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGCKVQ 180
Db 145 EEKKVPINLWDGKQTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGCKVQ 204
QY 181 RGLIVHSSSEGSTSVSYDLFDAQOYQPTDLLRIYRDNTTISSTSLSLSLYLYTT 233
Db 205 RGLIVHSSSEGSTSVSYDLFDAQOYQPTDLLRIYRDNKTINSENLDLYLYTT 257

RESULT 3
US-08-484-223B-112
; Sequence 112, Application US/08484223B
; Patent No. 6020459
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 245
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,223B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-484-223B-112

Query Match 90.9%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps

QY 1 SEKSEENKDLRKRSLOQTALGNLQKQIYYNSKAITSEKSADQFLNTLLFKGFFTG 60
Db 25 SEKSEENKDLRKRSLOQTALGNLQKQIYYNSKAITSEKSADQFLNTLLFKGFFTG 84

```


QY 61 HPYNDLLVDLGGTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
||||| : : : : :
Db 85 HPYNDLLVDLGGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
||||| : : : : :
Db 145 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSVSYDLFDAQGYPDTLLRIYRDNTTISSTLSLSLYLTT 233
||||| : : : : :
Db 205 RGLIVFHSSEGSVSYDLFDAQGYPDTLLRIYRDNTTINSENLDLHLYLTT 257

RESULT 4
US-08-919-597-112
; Sequence 112, Application US/08919597
; Patent No. 6054265
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; NUMBER OF SEQUENCES: 273
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/919,597
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/470,896
; FILING DATE: 06-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-020
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNTE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-919-597-112

Query Match 90.9%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;
QY 1 SEKSEINEKDLRKSELQGTALGNLKIYYYNKAITSEKSAOFLNTLLFKGFFTG 60
||||| : : : : :
Db 25 SEKSEINEKDLRKSELQARNLSNLRQIYYNKAITENKESDDQFLENTLLFKGFFTG 84

QY 61 HPYNDLLVDLGGTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
||||| : : : : :
Db 85 HPYNDLLVDLGGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
||||| : : : : :
Db 145 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSVSYDLFDAQGYPDTLLRIYRDNTTISSTLSLSLYLTT 233
||||| : : : : :
Db 205 RGLIVFHSSEGSVSYDLFDAQGYPDTLLRIYRDNTTINSENLDLHLYLTT 257

RESULT 5
US-08-475-668A-112
; Sequence 112, Application US/08475668A
; Patent No. 6060065
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/475,668A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-026
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNTE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-475-668A-112

Query Match 90.9%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;
QY 1 SEKSEINEKDLRKSELQGTALGNLKIYYYNKAITSEKSAOFLNTLLFKGFFTG 60
||||| : : : : :
Db 25 SEKSEINEKDLRKSELQARNLSNLRQIYYNKAITENKESDDQFLENTLLFKGFFTG 84
QY 61 HPYNDLLVDLGGTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
||||| : : : : :
Db 85 HPYNDLLVDLGGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
||||| : : : : :
Db 145 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSVSYDLFDAQGYPDTLLRIYRDNTTISSTLSLSLYLTT 233
||||| : : : : :
Db 205 RGLIVFHSSEGSVSYDLFDAQGYPDTLLRIYRDNTTINSENLDLHLYLTT 257

Db 145 EEKVPINLWIDGKQTPVIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTSLSLSLYLYTT 233
Db 205 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNKTINSENHLIDLILYLYTT 257

RESULT 6
US-08-485-551A-112
; Sequence 112, Application US/08485551A
; Patent No. 6068973
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.551A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-485-551A-112

Query Match 90.9%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELQGTALGNLKQIYYNYSKAITSEKSAQOFLNTLLFKGFFTG 60
Db 25 SEKSEINEKDLRKSELQRLNLSNLQIYYNYSKAITENKESDDQFLNTLLFKGFFTG 84
QY 61 HPWYNDLLVDLGSATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 85 HPWYNDLLVDLGSKATNRYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTPVIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
Db 145 EEKVPINLWIDGKQTPVIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTSLSLSLYLYTT 233
QY 181 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTSLSLSLYLYTT 233

Db 205 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNKTINSENHLIDLILYLYTT 257

RESULT 7
US-08-471-913A-112
; Sequence 112, Application US/08471913A
; Patent No. 6093794
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING EPSTEIN-BARR VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 214
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/471.913A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-471-913A-112

Query Match 90.9%; Score 1107; DB 3; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELQGTALGNLKQIYYNYSKAITSEKSAQOFLNTLLFKGFFTG 60
Db 25 SEKSEINEKDLRKSELQRLNLSNLQIYYNYSKAITENKESDDQFLNTLLFKGFFTG 84
QY 61 HPWYNDLLVDLGSATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 85 HPWYNDLLVDLGSKATNRYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTPVIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
Db 145 EEKVPINLWIDGKQTPVIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNNTTISSTSLSLSLYLYTT 233
Db 205 RGLIVFHSSEGSTVSVDLFDAGQYPTDLLRIYRDNKTINSENHLIDLILYLYTT 257

```
RESULT 8
US-08-485-264A-112
; Sequence 112, Application US/08485264A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING
; TITLE OF INVENTION: RESPIRATORY SYNCYTIAL VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485.264A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-021
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-485-264A-112

Query Match 90.9%; Score 1107; DB 4; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNKAITSSEKSADQFLNTLLFKGFTG 60
DB 25 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLNTLLFKGFTG 84
QY 61 HPWYNDLLVLDLSTAAATSEYEGSSVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLT 120
DB 85 HPWYNDLLVLDLSTAAATSEYEGSSVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
DB 145 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTT 233
DB 205 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTT 257

RESULT 9
US-08-474-349A-112
; Sequence 112, Application US/08474349A
; Patent No. 6333395
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING HUMAN PARAINFLUENZA
; TITLE OF INVENTION: VIRUS TRANSMISSION
; NUMBER OF SEQUENCES: 517
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/474.349A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-024
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 112:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-474-349A-112

Query Match 90.9%; Score 1107; DB 4; Length 257;
Best Local Similarity 89.7%; Pred. No. 1.6e-110;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNKAITSSEKSADQFLNTLLFKGFTG 60
DB 25 SEKSEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLNTLLFKGFTG 84
QY 61 HPWYNDLLVLDLSTAAATSEYEGSSVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLT 120
DB 85 HPWYNDLLVLDLSTAAATSEYEGSSVDLYGAYGYOCAGTGNKTCACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
DB 145 EEKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTT 233
DB 205 RGLIVFHSSEGSTVSVDLFDAGQGYPTDLLRIYRDNKTINSNLHIDLTYTT 257
```

APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
SUPERANTIGENS, AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005002
CURRENT APPLICATION NUMBER: US/08/896.933
EARLIER FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252.978
EARLIER FILING DATE: 1994-06-02
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 226
TYPE: PRT
ORGANISM: Staphylococcus aureas
US-08-896-933-24

Query Match 85.4%; Score 1040; DB 4; Length 226;
Best Local Similarity 85.2%; Pred. No. 2e-103;
Matches 196; Conservative 13; Mismatches 17; Indels 4; Gaps 1;
QY 4 SEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTGHPW 63
DB 1 SEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTGHPW 60
QY 64 YNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEK 123
DB 61 YNDLLVDKGSKDATNKYKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEE- 119
QY 124 KVPINLWDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGL 183
DB 120 ---VBKWDIGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGL 176
QY 184 IVFHSSEGSTVSYDLFDAQGQYPTDLLRIYRDNNTTISSTLSISLYLYTT 233
DB 177 IVFHSSEGSTVSYDLFDAQGQYPTDLLRIYRDNNTTINSENHIDIYLYTT 226

RESULT 11
US-09-314-235-24
Sequence 24, Application US/09314235
Patent No. 6338845
GENERAL INFORMATION:
APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
SUPERANTIGENS, AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005004
CURRENT APPLICATION NUMBER: US/09/314.235
CURRENT FILING DATE: 1999-05-18
EARLIER APPLICATION NUMBER: 08/896.933
EARLIER FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252.978
EARLIER FILING DATE: 1994-06-02
EARLIER APPLICATION NUMBER: 07/891.718
EARLIER FILING DATE: 1992-06-01
EARLIER APPLICATION NUMBER: US91/00342
EARLIER FILING DATE: 1991-01-17
EARLIER APPLICATION NUMBER: 07/466.577
EARLIER FILING DATE: 1990-01-17
EARLIER APPLICATION NUMBER: 07/416.530
EARLIER FILING DATE: 1989-10-03
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 24
LENGTH: 226
TYPE: PRT
ORGANISM: Staphylococcus aureas
US-09-314-235-24

Query Match 85.4%; Score 1040; DB 4; Length 226;
Best Local Similarity 85.2%; Pred. No. 2e-103;
Matches 196; Conservative 13; Mismatches 17; Indels 4; Gaps 1;
QY 4 SEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTGHPW 63

DB 1 SEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTGHPW 60
QY 64 YNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEEK 123
DB 61 YNDLLVDKGSKDATNKYKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEE- 119
QY 124 KVPINLWDGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGL 183
DB 120 ---VBKWDIGKQTTVPIDKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRGL 176
QY 184 IVFHSSEGSTVSYDLFDAQGQYPTDLLRIYRDNNTTISSTLSISLYLYTT 233
DB 177 IVFHSSEGSTVSYDLFDAQGQYPTDLLRIYRDNNTTINSENHIDIYLYTT 226

RESULT 12
US-08-486-099-113
Sequence 113, Application US/08486099
Patent No. 6013263
GENERAL INFORMATION:
APPLICANT: Bolognesi, Dani P.
APPLICANT: Matthews, Thomas J.
APPLICANT: Wild, Carl T.
APPLICANT: Barney, Shawn O.
APPLICANT: Lambert, Dennis M.
APPLICANT: Petteway, Stephen R.
APPLICANT: Langlois, Alphonse J.
TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HEPATITIS
TITLE OF INVENTION: B VIRUS TRANSMISSION
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486.099
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7872-031
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 257 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-486-099-113

Query Match 77.8%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 1.7e-93;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;
QY 1 SEKSEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTG 60
DB 25 SEKSEEINEKDLRKSELOQTALGNLKOIYYNKAITSSEKSADQFLTNLLFKGFTG 84

[illegible]

RESULT 13
US-08-360-107A-123
; Sequence 123, Application US/08360107A
; Patent No. 6017536
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR INHIBITION
; TITLE OF INVENTION: OF MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 149
; CORRESPONDENCE ADDRESS:

	Query Match	77.8%	Score 948;	DB 3;	Length 257;
	Best Local Similarity	76.4%;	Pred. No.	1.7e-93;	
	Matches	178;	Conservative	21;	Mismatches 34; Indels 0; Gaps 0;
Qy	1	SEKSEEINEKDRLKKSELQGTAALGNLKOIYYNYSKAITSSEKSADQFULTLLFKGFFTG	60		
Dd	25	SEKSEEINEKDRLKKSELQGTAALGNLKAIYYNEKAKTENKESHDPFLQHILFKGFPTD	84		
Qy	61	HPWYNDLLVLGGSTAATSEYEGSSVDLYGYAYYCQCAGGTENKTACMYGVVTLHDNNRKL	120		
Dd	85	HSWNNDLVDFDSKDIIVDKYKGKVLDLYGYAYYCQCAGGTENKTACMYGVVTLHDNNRKL	144		

121	QY	BEKKVPINLWDGQOTVPTDKVKTSKKEVTVQELDLQARHYLHGKFGILYNSDSFGKKVQ	180
145	Db	BEKKVPINLWDGQOTVPTDKVKTSKKEVTVQELDLQARHYLHGKFGILYNSDSFGKKVQ	204
181	QY	RGILVHSRSGSVSVYDLFPAQGGYPTDLLRIYRDNTTISSTLSLSLYLT	233
205	Db	RGILVHSTSPGVNVDLFGAOGGYNTLLRIYRDNTTINSENMHIDVLYTS	257

RESULT 14
 US-08-484-223B-113
 ; Sequence 113, Application US/08484223B
 ; Patent No. 6020459
 ; GENERAL INFORMATION:
 ; APPLICANT: Bognesi, Dani P.
 ; APPLICANT: Matthews, Thomas J.
 ; APPLICANT: Wild, Carl T.
 ; APPLICANT: Barney, Shawn O.
 ; APPLICANT: Lambert, Dennis M.
 ; APPLICANT: Petteway, Stephen R.
 ; APPLICANT: Langlois, Alphonse J.
 ; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF
 ; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV
 ; TITLE OF INVENTION: TRANSMISSION
 ; NUMBER OF SEQUENCES: 245
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,223B
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Coruzzi, Laura A.
 ; REGISTRATION NUMBER: 30,742
 ; REFERENCE/DOCKET NUMBER: 7872-029
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 113:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 257 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS:
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: protein
 ; US-08-484-223B-113

Query Match	77.8%	Score 948;	DB 3;	Length 257;
Best Local Similarity	76.4%;	Pred. No. 1.7e-93;		
Matches	178;	Conservative 21;	Mismatches 34;	Indels 0; Gaps 0;
Qy	1	SEKSEINEKDLRKKSELOQTALGNLQIYYIYNKAITSEKSAQDFLTNTLLKGFFTG	60	
Db	25	SEKSEINEKDLRKKSELOQTALGNLQIYYIYNKAKTENKESHQDFLQHTLLKGFPTD	84	
Qy	61	HPWYNLLVLGSTAATSEYEGSSVDLYGAYGYQCAGGTENKTACMYGGVTLHDNNRLT	120	
Db	85	HSWYNLLLVDFDSKDIYDKYKGVKVDLYGAYGYQCAGGTENKTACMYGGVTLHDNNRLT	144	
Qy	121	EKKVPINLWIDGKQTTPIDKVKTSKKEVTQBELDQLQARHLYLHGKFGLYNSDSFGGRKVQ	180	


```
RESULT 17
US-08-485-551A-113
; Sequence 113, Application US/08485551A
; Patent No. 6068973
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
; APPLICANT: Matthews, Thomas J.
; APPLICANT: Wild, Carl T.
; APPLICANT: Barney, Shawn O.
; APPLICANT: Lambert, Dennis M.
; APPLICANT: Petteway, Stephen R.
; APPLICANT: Langlois, Alphonse J.
; TITLE OF INVENTION: METHODS FOR INHIBITION OF MEMBRANE
; TITLE OF INVENTION: FUSION-ASSOCIATED EVENTS, INCLUDING INFLUENZA VIRUS
; TITLE OF INVENTION: TRANSMISSION
; NUMBER OF SEQUENCES: 211
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,551A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7872-023
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-485-551A-113

Query Match 77.8%; Score 948; DB 3; Length 257;
Best Local Similarity 76.4%; Pred. No. 1.7e-93;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNSKAITSEKSADOFLLNTLLFKGFFTG 60
Db 25 SEKSEINEKDLRKSELOGTALGNLKOIYYNEKAKTENKESHQDQLQHTILFKGFFTD 84
Qy 61 HPWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 85 HSWYNLLVDFDSKDIDVKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
Qy 121 EEKKVPINLWIDGKQTPVIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKVKQ 180
Db 145 EEKKVPINLWLDGKQNTVPLETVKTNKKNVTVOELDLQARRYLQEKYLNLYNSDVFQKVKQ 204
Qy 181 RGLIVPHSEGSTSVSYDLFDAOGQYPTDLLRIYRONTTISSTLSLSLYTT 233
Db 205 RGLIVPHTSTEPSVNYDLFGAOGQYNTLLRIYRONTKNTINSENMHIDIYLYTS 257

RESULT 18
US-08-471-913A-113
; Sequence 113, Application US/08471913A
; Patent No. 6228983
; GENERAL INFORMATION:
; APPLICANT: Bolognesi, Dani P.
```


CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/446,918A
FILING DATE: 18-MAY-1995
CLASSIFICATION: 552
ATTORNEY/AGENT INFORMATION:
NAME: Kovarik, Joseph E.
REGISTRATION NUMBER: 33,005
REFERENCE/DOCKET NUMBER: 2879-29
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-446-918A-4

Query Match 77.5%; Score 944; DB 1; Length 233;
Best Local Similarity 76.3%; Pred. No. 4e-93;
Matches 177; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 2 EKSEINEKDLRKSELOGTALGNLKOIYYYNKAITTSSEKSADQFLTNLLFKGFFTGH 61
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 2 EKSEINEKDLRKSELOGTALGNLKOIYYYNKAKTENKESHDOFLQHTILFKGFFTDH 61
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
QY 62 PWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 62 SWYNLLVDFDSKIDVYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
QY 122 EKKVPINLWIDGKQTPVDIKVYKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQVOR 181
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 122 EKKVPINLWIDGKQTPVLETVKNNVTVOELDLQARRYLOEKYLYNSDVFQKQVOR 181
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
QY 182 GLIVFHSSSGSVSYDLFDAQOQYPTLLRIYRDNTTISSTLSISLYLYTT 233
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 182 GLIVFHTSTPSVNYDLFGAQOYSNTLLRIYRDNTINSENHDIYLYTS 233
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :

RESULT 22
US-08-580-806-4
Sequence 4, Application US/08580806
Patent No. 5935568
GENERAL INFORMATION:
APPLICANT: Dow, Steve W.
APPLICANT: Elmslie, Robyn E.
APPLICANT: Potter, Terence A.
TITLE OF INVENTION: GENE THERAPY FOR EFFECTOR CELL REGULATION
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A.
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/580,806

FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2879-29-C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 233 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-580-806-4

Query Match 77.5%; Score 944; DB 2; Length 233;
Best Local Similarity 76.3%; Pred. No. 4e-93;
Matches 177; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

QY 2 EKSEINEKDLRKSELOGTALGNLKOIYYYNKAITTSSEKSADQFLTNLLFKGFFTGH 61
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 2 EKSEINEKDLRKSELOGTALGNLKOIYYYNKAKTENKESHDOFLQHTILFKGFFTDH 61
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
QY 62 PWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 62 SWYNLLVDFDSKIDVYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 121
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
QY 122 EKKVPINLWIDGKQTPVDIKVYKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQVOR 181
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 122 EKKVPINLWIDGKQTPVLETVKNNVTVOELDLQARRYLOEKYLYNSDVFQKQVOR 181
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
QY 182 GLIVFHSSSGSVSYDLFDAQOQYPTLLRIYRDNTTISSTLSISLYLYTT 233
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 182 GLIVFHTSTPSVNYDLFGAQOYSNTLLRIYRDNTINSENHDIYLYTS 233
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :

RESULT 23
US-08-896-933-23
Sequence 23, Application US/08896933
Patent No. 6221351
GENERAL INFORMATION:
APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
SUPERANTIGENS, AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005002
CURRENT APPLICATION NUMBER: US/08/896,933
CURRENT FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252,978
EARLIER FILING DATE: 1994-06-02
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 23
LENGTH: 232
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-896-933-23

Query Match 76.5%; Score 931.5; DB 4; Length 232;
Best Local Similarity 76.0%; Pred. No. 8.7e-92;
Matches 177; Conservative 21; Mismatches 34; Indels 1; Gaps 1;

QY 1 SEKSEINEKDLRKSELOGTALGNLKOIYYYNKAITTSSEKSADQFLTNLLFKGFFTG 60
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 1 SEKSEINEKDLRKSELOGTALGNLKOIYYYNKAKTENKESHDOFL-HTILFKGFFTD 59
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
QY 61 HPWYNLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 120
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
Db 60 HSWYNLLVDFDSKIDVYKKGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTE 119
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :
QY 121 EKKVPINLWIDGKQTPVDIKVYKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQVOR 180
||||| :||||| :||||| :||||| :||||| :||||| :||||| :||||| :

```
Db 120 EKKVPINLWLDGKONTVPLETVKTNKNTVQELDPQARRYLQEKYNLYNSDVFQKVQ 179
QY 181 RGLIVFHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNTTISSTLSISLYLYTT 233
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|
Db 180 RGLIVFHTSTEPSVNDLFGAQOQYNTLLRIYRDNTKNTINSENMMHIDIYLYTS 232
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|

RESULT 24
US-09-314-235-23
; Sequence 23, Application US/09314235
; Patent No. 633845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; CURRENT FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416,530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-23

Query Match 76.5%; Score 931.5; DB 4; Length 232;
Best Local Similarity 76.0%; Pred. No. 8.7e-92;
Matches 177; Conservative 21; Mismatches 34; Indels 1; Gaps 1;

QY 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTG 60
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|
Db 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTG 59
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|

QY 61 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGTGNKTCMYGGVTLHDNNRLT 120
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|
Db 60 HSWYNDLLVDFDSKDIVDKYKGYDLYGAYGYOCAGTGNKTCMYGGVTLHDNNRLT 119
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|

QY 121 EKKVPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHLYHGKFGLYNSDSFGKVQ 180
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|
Db 120 EKKVPINLWLDGKONTVPLETVKTNKNTVQELDPQARRYLQEKYNLYNSDVFQKVQ 179
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|

QY 181 RGLIVFHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNTTISSTLSISLYLYTT 233
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|
Db 180 RGLIVFHTSTEPSVNDLFGAQOQYNTLLRIYRDNTKNTINSENMMHIDIYLYTS 232
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|

RESULT 25
US-09-144-776B-2
; Sequence 2, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
; Mark A. Olson
; Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
; Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army MRC -504 Scott Street
```

```
; MCMR-JA (Charles H. Harris-Patent
; Atty)
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,776B
; FILING DATE: 01-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/882,431
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 257
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-144-776B-2

Query Match 75.9%; Score 925; DB 4; Length 257;
Best Local Similarity 75.1%; Pred. No. 5e-91;
Matches 175; Conservative 21; Mismatches 37; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTG 60
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|
Db 25 SEKSEINEKDLRKSELQGTALGNLKOIYYNKAITSSEKSADQFLTNTLLFKGFTG 84
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|

QY 61 HPWYNDLLVLDGSTAATSEYEGSSVDLYGAYGYOCAGTGNKTCMYGGVTLHDNNRLT 120
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|
Db 85 HSWYNDLLVDFDSKDIVDKYKGYDLYGAYGYOCAGTGNKTCMYGGVTLHDNNRLT 144
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|

QY 121 EKKVPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHLYHGKFGLYNSDSFGKVQ 180
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|
Db 145 EKKVPINLWLDGKONTVPLETVKTNKNTVQELDLQARRYLQEKYNLYNSDVFQKVQ 204
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|

QY 181 RGLIVFHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNTTISSTLSISLYLYTT 233
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|
Db 205 RGLIVFHTSTEPSVNDLFGAQOQYNTLLRIYRDNTKNTINSENMMHIDIYLYTS 257
|||||:| :|||:| |||||:| |||||:| ||:| :| :|||:|

RESULT 26
US-09-144-776B-4
; Sequence 4, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
; Mark A. Olson
; Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
; Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army MRC -504 Scott Street
; MCMR-JA (Charles H. Harris-Patent
; Atty)
; CITY: FORT DETRICK
```

STATE: MARYLAND
COUNTRY: USA
ZIP: 21702-5012
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,776B
FILING DATE: 01-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/882,431
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Charles H. Harris
REGISTRATION NUMBER: 34,616
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 619-2065
TELEFAX: (301) 619-7714
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 233
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-144-776B-4

Query Match 75.6%; Score 921; DB 4; Length 233;
Best Local Similarity 75.0%; Pred. No. 1.2e-90;
Matches 174; Conservative 21; Mismatches 37; Indels 0; Gaps 0;
QY 2 EKSEENKDLRKSELOGTALGNLQIYYNSKAITSSSEKSDAQFLTNLLFKGFFTG 61
Db 2 EKSEENKDLRKSELOGTALGNLQIYYNEKAKTENKESHDPQHTILFKGFFTDH 61
QY 62 PWNLDVLGSAATSEYEGSSVDLYGAYGYOCAGGTPNKTCACMYGGVTLHDNNRLTE 121
Db 62 SWNDLLVRFDSKDIYDKYKGVLDLYGAYGYOCAGGTPNKTCACMYGGVTLHDNNRLTE 121
QY 122 EKKVPINLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQR 181
Db 122 EKKVPINLWLDGKONTVPLETVTKTKNNVTVOELDLQARRYLQEKYLNYSNDSVDFGKVQR 181
QY 182 GLIVFHSSEGSTVSYDLFDAOGQYPTDLLRIYRDNTTISSTLSISLYTT 233
Db 182 GLIVFHTSTEPSVNYDLFQAQGYSNLTLRIYRDNKTINSENHIDIYLYTS 233

RESULT 27
US-08-896-933-25
Sequence 25, Application US/08896933
Patent No. 6221351
GENERAL INFORMATION:
APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
SUPERANTIGENS, AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005002
CURRENT APPLICATION NUMBER: US/08/896,933
EARLIER FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252,978
EARLIER FILING DATE: 1994-06-02
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 25
LENGTH: 228
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-08-896-933-25

Query Match 49.7%; Score 605; DB 4; Length 228;
Best Local Similarity 52.0%; Pred. No. 7.9e-57;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;
QY 7 INEKDLRKSELOGTALGNLQIYYNSKAITSSSEKSDAQFLTNLLFKGFFTGHPWYND 66
Db 2 VKEKELHKSELSSTALNNMKHSYADKNPTIGENKSTGDOFLENTLLYKFFTDLINFED 61
QY 67 LLVDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTCACMYGGVTLHDNNRLTEKKVP 126
Db 62 LLINFNSKEMAQHFKSKNVDVPIRISINCYGGEIDRTACTYGGVTPHEGNKLKERKKIP 121
QY 127 INLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRLIVF 186
Db 122 INLWINGVQKEVSLDKVQTDKKNVTVOELDAQARRYLQDKLYNNDTLGGKIQRGKIEF 181
QY 187 HSSEGSTVSYDLFDAOGQYPTDLLRIYRDNTTISSTLSISLYLY 231
Db 182 DSSDGSKVSYDLFDVKGDFPEKQRIYSDNKTLSLSTELHIDIYLY 226

RESULT 28
US-09-314-235-25
Sequence 25, Application US/09314235
Patent No. 6338845
GENERAL INFORMATION:
APPLICANT: Terman, David S.
TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
SUPERANTIGENS, AND RELATED COMPOUNDS
FILE REFERENCE: 09629/005004
CURRENT APPLICATION NUMBER: US/09/314,235
EARLIER FILING DATE: 1999-05-18
EARLIER APPLICATION NUMBER: 08/896,933
EARLIER FILING DATE: 1997-07-18
EARLIER APPLICATION NUMBER: 08/252,978
EARLIER FILING DATE: 1994-06-02
EARLIER APPLICATION NUMBER: 07/891,718
EARLIER FILING DATE: 1992-06-01
EARLIER APPLICATION NUMBER: US91/00342
EARLIER FILING DATE: 1991-01-17
EARLIER APPLICATION NUMBER: 07/466,577
EARLIER FILING DATE: 1990-01-17
EARLIER APPLICATION NUMBER: 07/416,530
EARLIER FILING DATE: 1989-10-03
NUMBER OF SEQ ID NOS: 34
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 25
LENGTH: 228
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-314-235-25

Query Match 49.7%; Score 605; DB 4; Length 228;
Best Local Similarity 52.0%; Pred. No. 7.9e-57;
Matches 117; Conservative 33; Mismatches 75; Indels 0; Gaps 0;
QY 7 INEKDLRKSELOGTALGNLQIYYNSKAITSSSEKSDAQFLTNLLFKGFFTGHPWYND 66
Db 2 VKEKELHKSELSSTALNNMKHSYADKNPTIGENKSTGDOFLENTLLYKFFTDLINFED 61
QY 67 LLVDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTCACMYGGVTLHDNNRLTEKKVP 126
Db 62 LLINFNSKEMAQHFKSKNVDVPIRISINCYGGEIDRTACTYGGVTPHEGNKLKERKKIP 121
QY 127 INLWIDGKOTTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQRLIVF 186
Db 122 INLWINGVQKEVSLDKVQTDKKNVTVOELDAQARRYLQDKLYNNDTLGGKIQRGKIEF 181
QY 187 HSSEGSTVSYDLFDAOGQYPTDLLRIYRDNTTISSTLSISLYLY 231
Db 182 DSSDGSKVSYDLFDVKGDFPEKQRIYSDNKTLSLSTELHIDIYLY 226

```
RESULT 29
US-09-144-776B-19
; Sequence 19, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
;           Mark A. Olson
;           Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
;           Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army MRC -504 Scott Street
;           MCMR-JA (Charles H. Harris-Patent
;           Atty)
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,776B
; FILING DATE: 01-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/882,431
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 82
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-144-776B-19
Query Match 33.3%; Score 406; DB 4; Length 82;
Best Local Similarity 89.0%; Pred. No. 3.4e-36;
Matches 73; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 43 SADOFLNTLLFKGFTGHPWYNDLLVLDLGSSTAATSEYEGSSVDLYGAYGYQCAGGTPN 102
| ||||| :||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 SDOQFLNTLLFKGFTGHPWYNDLLVLDLGSKDATNKYKGVKVDLYGAYGYQCAGGTPN 60
QY 103 KTACMYGGVTLHDNNRLTEKK 124
| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 KTACMYGGVTLHDNNRLTEKK 82
RESULT 30
US-09-144-776B-17
; Sequence 17, Application US/09144776B
; Patent No. 6399332
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
;           Mark A. Olson
;           Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
;           Vaccines
```

```
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles H. Harris
STREET: US Army MRC -504 Scott Street
        MCMR-JA (Charles H. Harris-Patent
        Atty)
CITY: FORT DETRICK
STATE: MARYLAND
COUNTRY: USA
ZIP: 21702-5012
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,776B
FILING DATE: 01-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/882,431
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Charles H. Harris
REGISTRATION NUMBER: 34,616
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 619-2065
TELEFAX: (301) 619-7714
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 82
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 17:
US-09-144-776B-17
Query Match 29.7%; Score 362; DB 4; Length 82;
Best Local Similarity 79.3%; Pred. No. 1.8e-31;
Matches 65; Conservative 4; Mismatches 13; Indels 0; Gaps 0;
QY 43 SADOFLNTLLFKGFTGHPWYNDLLVLDLGSSTAATSEYEGSSVDLYGAYGYQCAGGTPN 102
| ||||| :||:||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1 SHDQFLQHTLILFKGFTDHSWYNDLLVDFDSKDIVDKYGKVDLYGAYGYQCAGGTPN 60
QY 103 KTACMYGGVTLHDNNRLTEKK 124
| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 KTACMYGGVTLHDNNRLTEKK 82
RESULT 31
US-08-896-933-28
; Sequence 28, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896,933
; CURRENT FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-28
```

```
Query Match      22.8%; Score 277.5; DB 4; Length 238;
Best Local Similarity 32.5%; Pred. No. 1e-21;
Matches 76; Conservative 44; Mismatches 95; Indels 19; Gaps 9;

QY 11 DLKSKSELOGTALGNLKQIYYNSKAITSS--KSADQFLTNTLLFKGFTGHPWYNDLLV 69
Db 9 ELHKSSEFTGT-MGNKK--YLYDDHVSATKNSVDKFLAHLDIYINISDKLKNYDKVKT 65
QY 70 DLGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKATCMYGGVTLHDNNRLTEE 122
Db 66 ELLNEDLAKKYKDEVDVYGSNYVNCYFSSKDNVKGVTGGKTCMVGGITKHEGNHFDNG 125
QY 123 --KKVPINLWIDCKQTPIDPKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 180
Db 126 NLQNVLRIVY-ENKRTISFE-VQTDKSVTAQELDIKARNFINKKNLYEFNS--SPYE 181
QY 181 RGLIVFHSSEGSVSVDLFDAGQYDP--TLRLIYFDNTTISSTLSISLYLT 232
Db 182 TGYIKFIENNGNTFWDLAPAPGDKFDQSKYLMYNDNKTVDKSKVKIEVHLTT 235

RESULT 32
US-09-314-235-28
; Sequence 28, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; EARLIER FILING DATE: 1999-05-18
; EARLIER FILING DATE: 1997-07-18
; EARLIER FILING DATE: 1994-06-02
; EARLIER FILING DATE: 1992-06-01
; EARLIER FILING DATE: 1991-01-17
; EARLIER FILING DATE: 1990-01-17
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28
; LENGTH: 238
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-28

Query Match      22.8%; Score 277.5; DB 4; Length 238;
Best Local Similarity 32.5%; Pred. No. 1e-21;
Matches 76; Conservative 44; Mismatches 95; Indels 19; Gaps 9;

QY 11 DLKSKSELOGTALGNLKQIYYNSKAITSS--KSADQFLTNTLLFKGFTGHPWYNDLLV 69
Db 9 ELHKSSEFTGT-MGNKK--YLYDDHVSATKNSVDKFLAHLDIYINISDKLKNYDKVKT 65
QY 70 DLGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKATCMYGGVTLHDNNRLTEE 122
Db 66 ELLNEDLAKKYKDEVDVYGSNYVNCYFSSKDNVKGVTGGKTCMVGGITKHEGNHFDNG 125
QY 123 --KKVPINLWIDCKQTPIDPKVTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKQV 180
Db 126 NLQNVLRIVY-ENKRTISFE-VQTDKSVTAQELDIKARNFINKKNLYEFNS--SPYE 181
QY 181 RGLIVFHSSEGSVSVDLFDAGQYDP--TLRLIYFDNTTISSTLSISLYLT 232
Db 182 TGYIKFIENNGNTFWDLAPAPGDKFDQSKYLMYNDNKTVDKSKVKIEVHLTT 235
```

```
RESULT 33
US-08-896-933-29
; Sequence 29, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896,933
; CURRENT FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-08-896-933-29

Query Match      22.2%; Score 270.5; DB 4; Length 221;
Best Local Similarity 33.8%; Pred. No. 5.1e-21;
Matches 76; Conservative 42; Mismatches 88; Indels 19; Gaps 10;

QY 16 SELQGTAL-GNLKQIYY-YNSKAIT-SSEKSAQDFLTNTLLFKGFTGHPWYNDLLV 72
Db 7 SQLHSSLVKNLQNIYFLYEGDPVTHENVKSVQDQLLHSHDLIYN--VSGPNYDKLATELK 63
QY 73 STAATSEYEGSSVDLYGAYGYQC-AGGTPNKATCMYGGVTLHDNNRLTEEKVPINLWI 131
Db 64 NOEMATLFDKDNVDIYGVVEYHLCYLCENASACIYGGVTNHEGHNHLEIPKKIYVYK 123
QY 132 DGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKQVGLIVFHS 191
Db 124 DGIQ-SLSFD-IETNKKMVTQAQELDYKRYLTDNKLTYNGP--SKYETGYIKFIPKN 179
QY 192 STVSYDLFD----AQQYPTDTLRIYRDNNTTISSTLSISLYLT 232
Db 180 ESFWDFLFPPEPTOSKY---LMIYKDNETLDSNTSQIEVLYLT 220

RESULT 34
US-09-314-235-29
; Sequence 29, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416,530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 29
; LENGTH: 221
; TYPE: PRT
; ORGANISM: Streptococcus pyogenes
US-09-314-235-29
```

```
Query Match      22.2%; Score 270.5; DB 4; Length 221;
Best Local Similarity 33.8%; Pred. No. 5.1e-21;
Matches 76; Conservative 42; Mismatches 28; Indels 19; Gaps 10;

Qy 16 SELQGTAL-GNLKQIYV-VNSKAIT--SSPKSADQFLTNLLFKGFFTHGPWYNLDLVLG 72
Db 7 SOLHRSSLVKNLQNIYFLYEGDPVTHENVKSVQDLSSHDLIYN---VSGPNYDKLKTCLK 63

Qy 73 STAATSEYEGSSVDLYGAYGYQC-AGGTPNKACMYGGVTLHDNNRLTEKKVPINLWI 131
Db 64 NQEMATLPKDKNVDIYGVYEHLCVLCNAERSACIYGGVTNHEGNHLEIPKKIVKYSI 123

Qy 132 DSKQTTVPIDKYKTSKEVTVQELDLQARHYLHGKGLYNSDFGKVGKVGORGLIVFHSSEG 191
Db 124 DGIQ-SLSFD-IETNKKMVTAGQELDYKVRKYLTDNKLQVLTNGP--SKYETGYIKFIPK 179

Qy 192 STVSYDLDF---AQOXPDTLLRIYRDNTTISSTLSLSLYLT 232
Db 180 ESEWDLFPPEFTQSKY----LMYKDNELDSNTSQIEVILTT 220

RESULT 35
US-08-446-918A-2
; Sequence 2, Application US/08446918A
; Patent No. 5705151
; GENERAL INFORMATION:
; APPLICANT: Dow, Steve W.
; APPLICANT: Elmslie, Robyn E.
; TITLE OF INVENTION: GENE THERAPY FOR T CELL REGULATION
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446.918A
; FILING DATE: 18-MAY-1995
; CLASSIFICATION: 552
; ATTORNEY/AGENT INFORMATION:
; NAME: Kovarik, Joseph E.
; REGISTRATION NUMBER: 33,005
; REFERENCE/DOCKET NUMBER: 2879-29
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-446-918A-2

Query Match      22.0%; Score 267.5; DB 1; Length 255;
Best Local Similarity 31.7%; Pred. No. 1.3e-20;
Matches 77; Conservative 43; Mismatches 104; Indels 19; Gaps 8;

Qy 2 EKSEINEKDLRKKSLOGTALGNLKOIYYNSKAITSEKSDAQFLNTLLFKGFFTHG 61
Db 17 ESQDPKPKDELHKSKFTG-LMENMK-VLYDDNHVSAINKSIDQFLYFDLIYSKDTKL 74

Qy 62 PWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKACMYGGV 111
Db 75 GNYDNRVFEKNKDLADKYDKYVDVFGANYYYQCYFSKKTNDINSHQTDKRTKCMYGGV 134

Qy 112 TLHDNNRLTEKKVPINLWIDGKQTTVPIDKYKTSKEVTVQELDLQARHYLHGKGLYN 171
Db 135 TEHNGNQLDKYRSITVRVFEDEGK-NLLSFD-VQTNKKKVTAGQELDYLTRHYLVKRNKKLYE 192

Qy 172 SDSFGKVGKVGORGLIVFHSSEGSPVSYDLFDAOQOYDP--TLLRIYRDNTTISSTLSLSI 229
Db 193 FNN--SPIETGYIKFIENENS--FWYDMPAPGDKFDQSKYLMYNDNMKVDKVKIEVY 249
```

```
Qy 112 TLHDNNRLTEKKVPINLWIDGKQTTVPIDKYKTSKEVTVQELDLQARHYLHGKGLYN 171
Db 135 TEHNGNQLDKYRSITVRVFEDEGK-NLLSFD-VQTNKKKVTAGQELDYLTRHYLVKRNKKLYE 192

Qy 172 SDSFGKVGKVGORGLIVFHSSEGSPVSYDLFDAOQOYDP--TLLRIYRDNTTISSTLSLSI 229
Db 193 FNN--SPIETGYIKFIENENS--FWYDMPAPGDKFDQSKYLMYNDNMKVDKVKIEVY 249

Qy 230 LYT 232
Db 250 LTT 252

RESULT 36
US-08-580-806-2
; Sequence 2, Application US/08580806
; Patent No. 5935568
; GENERAL INFORMATION:
; APPLICANT: Dow, Steve W.
; APPLICANT: Elmslie, Robyn E.
; APPLICANT: Potter, Terence A.
; TITLE OF INVENTION: GENE THERAPY FOR EFFECTOR CELL REGULATION
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/580.806
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Connell, Gary J.
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER: 2879-29-C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 255 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-580-806-2

Query Match      22.0%; Score 267.5; DB 2; Length 255;
Best Local Similarity 31.7%; Pred. No. 1.3e-20;
Matches 77; Conservative 43; Mismatches 104; Indels 19; Gaps 8;

Qy 2 EKSEINEKDLRKKSLOGTALGNLKOIYYNSKAITSEKSDAQFLNTLLFKGFFTHG 61
Db 17 ESQDPKPKDELHKSKFTG-LMENMK-VLYDDNHVSAINKSIDQFLYFDLIYSKDTKL 74

Qy 62 PWYNLLVDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKACMYGGV 111
Db 75 GNYDNRVFEKNKDLADKYDKYVDVFGANYYYQCYFSKKTNDINSHQTDKRTKCMYGGV 134

Qy 112 TLHDNNRLTEKKVPINLWIDGKQTTVPIDKYKTSKEVTVQELDLQARHYLHGKGLYN 171
Db 135 TEHNGNQLDKYRSITVRVFEDEGK-NLLSFD-VQTNKKKVTAGQELDYLTRHYLVKRNKKLYE 192

Qy 172 SDSFGKVGKVGORGLIVFHSSEGSPVSYDLFDAOQOYDP--TLLRIYRDNTTISSTLSLSI 229
Db 193 FNN--SPIETGYIKFIENENS--FWYDMPAPGDKFDQSKYLMYNDNMKVDKVKIEVY 249
```


TELEFAX: (301) 619-7714
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 251
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 16:
US-09-144-776B-16

Query Match 21.4%; Score 263.5; DB 4; Length 251;
Best Local Similarity 32.1%; Pred. No. 3.5e-20;
Matches 76; Conservative 45; Mismatches 97; Indels 19; Gaps 10;
Qy 4 SEEINEKDLRKSELOGTALGNLQIYY-YNSKAIT-SSEKSADQFLTNLLFKGFTGH 60
Db 25 SOEVAQDPDPQSQRHSSLVKNLQNIYFLYEGDVTTHENKVSVDQLRSHDLIYN---VS 81
Qy 61 HPWYNDLLVLDGTAATSEYEGSSVDLYGAYGYQC-AGGTPNKTCACMYGGVTLHNNRL 119
Db 82 GPYDKLTKLKNQEMATLFKDKNDYIGVEYHLCYLCENAESACIYGGVTHHEGNHL 141
Qy 120 TEEKKVPINLWDGKQTTPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFGKV 179
Db 142 EIPKKIVVKSIDGIQ-SLSFD-IETNKKMVTQAELDKVKRYLTDNKLQLYTNGP--SKY 197
Qy 180 QRLGVHFSSEGSTVSYDLFD---AQQYDPDTPLLRIYRDNTTISSTLSLSLYLT 232
Db 198 ETGYIKFIPKNKESFWDFPEPEPTQSKY----LMIYKDNELDSNTSQIEVYLTT 250

RESULT 40
US-08-896-933-26
; Sequence 26, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896.933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252.978
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-26

Query Match 21.4%; Score 260.5; DB 4; Length 239;
Best Local Similarity 32.0%; Pred. No. 6.8e-20;
Matches 77; Conservative 41; Mismatches 104; Indels 19; Gaps 8;
Qy 2 EKSEINEKDLRKSELOGTALGNLQIYYNSKAITSSKSADQFLTNLLFKGFTGH 61
Db 1 ESQDPKDELHKSSKFTG-LMENMK-VLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKL 58
Qy 62 PWYNDLLVLDGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTCACMYGGV 111
Db 59 GNYDNVRVEFNKDLADYKDKYVDVFGANYYYQCYFSKKTNDINSHQTDKRTKTCMYGGV 118
Qy 112 TLHNNRLTEKKVPIINLWIDGKQTTPIDKVKTSKEVTVOELDLQARHYLHGKFLY 171
Db 119 TEHNGNQLDKYRSLTVRVFEDGK-NLLSFD-VQTNKKKVTQAELDYLTRHYLVKNKKLYE 176
Qy 172 SDSFGKVGQRLIVFHSSEGSTVSYDLFDAGQGYPD--TLRIYRDNTTISSTLSISLY 229
Db 177 FNN--SPYETGYIKFIENENS-FWYDLMPAPGDKFDQSKYLMNNDNMVDSKDVKIEVY 233

Qy 230 L 230
Db 234 L 234

RESULT 41
US-09-314-235-26
; Sequence 26, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314.235
; CURRENT FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 08/896.933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252.978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891.718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466.577
; EARLIER FILING DATE: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416.530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-26

Query Match 21.4%; Score 260.5; DB 4; Length 239;
Best Local Similarity 32.0%; Pred. No. 6.8e-20;
Matches 77; Conservative 41; Mismatches 104; Indels 19; Gaps 8;
Qy 2 EKSEINEKDLRKSELOGTALGNLQIYYNSKAITSSKSADQFLTNLLFKGFTGH 61
Db 1 ESQDPKDELHKSSKFTG-LMENMK-VLYDDNHVSAINVKSIDQFLYFDLIYSIKDTKL 58
Qy 62 PWYNDLLVLDGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTCACMYGGV 111
Db 59 GNYDNVRVEFNKDLADYKDKYVDVFGANYYYQCYFSKKTNDINSHQTDKRTKTCMYGGV 118
Qy 112 TLHNNRLTEKKVPIINLWIDGKQTTPIDKVKTSKEVTVOELDLQARHYLHGKFLY 171
Db 119 TEHNGNQLDKYRSLTVRVFEDGK-NLLSFD-VQTNKKKVTQAELDYLTRHYLVKNKKLYE 176
Qy 172 SDSFGKVGQRLIVFHSSEGSTVSYDLFDAGQGYPD--TLRIYRDNTTISSTLSISLY 229
Db 177 FNN--SPYETGYIKFIENENS-FWYDLMPAPGDKFDQSKYLMNNDNMVDSKDVKIEVY 233
Qy 230 L 230
Db 234 L 234

RESULT 42
US-09-414-276-8
; Sequence 8, Application US/09414276
; Patent No. 6392121
; GENERAL INFORMATION:
; APPLICANT: Mason, Hugh
; APPLICANT: Palmer, Kenneth
; APPLICANT: Hefferon, Kathleen
; APPLICANT: Mor. Tsafir
; APPLICANT: Arntzen, Charles
; TITLE OF INVENTION: Gemini Virus Vectors for Gene Expression in Plants
; FILE REFERENCE: 4868/84453


```

; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 266
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-144-776B-6

```

Query Match	21.4%;	Score 260.5;	DB 4;	Length 256;
Best Local Similarity	30.7%;	Pred. No. 8e-20;		
Matches 75;	Conservative 43;	Mismatches 107;	Indels 19;	Gaps 7;

[illegible]

RESULT 44
US-09-144-776B-10
; Sequence 10, Application US/09144776B
; Patent No. 639932
; GENERAL INFORMATION:
; APPLICANT: Robert G. Ulrich,
; Mark A. Olson
; Sina Bavari
; TITLE OF INVENTION: Bacterial Superantigen
; Vaccines
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles H. Harris
; STREET: US Army NRCM - 504 Scott St
; MCMR-JA (Charles H. Harris
; Att'y)
; CITY: FORT DETRICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh 7.5
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/144,776B
; FILING DATE: 01-Sep-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/882,431
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Charles H. Harris
; REGISTRATION NUMBER: 34,616
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:

TELEPHONE: (301) 619-2065
TELEFAX: (301) 619-7714
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 239
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-144-776B-10

Query Match 21.0%; Score 255.5; DB 4; Length 239;
Best Local Similarity 32.1%; Pred. No. 2.3e-19;
Matches 75; Conservative 41; Mismatches 99; Indels 19; Gaps 8;
QY 11 DLRRKSELQGTALGNLKOIYYNKAITSSEKSAQOFLNTLLFKGFTGHPWYNDLLVD 70
DB 10 ELHKSSKFTG-LMENMK-VLYDDNHVSAINVAKSIDQFRYFDLIYSIKDTKLGNYDNVRVE 67
QY 71 LGSTAATSEYEGSSVDLYGAYGYQCAGG-----TPNKTACMYGGVTLHDNNRLT 120
DB 68 FKNKLADKYDKYVDVFGANAYQCAFSSKKTNDINSHQTDKRTCMYGGVTEHNGNQLD 127
QY 121 EEKYPINLWIDGKQTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVQ 180
DB 128 KYRSITVRVEEDGK-NLLSFD-VQTNKKKVTQAQELDYLTRHYLVKNKKLYEFNN--SPYE 183
QY 181 RGLIVFHSSEGTVDYDLFDAQGOYPD--TLRLYRDNTTISSTLSLSLYLT 232
DB 184 TGYIRFIENENS-FWYDMMPAPGDFQDSKYLMMYNDNKNVDSKDVKIEVLYLT 236

RESULT 45
US-09-144-776B-14
Sequence 14, Application US/09144776B
Patent No. 6399332
GENERAL INFORMATION:
APPLICANT: Robert G. Ulrich,
Mark A. Olson
Sina Bavari
TITLE OF INVENTION: Bacterial Superantigen
Vaccines
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles H. Harris
STREET: US Army MRC -504 Scott Street
MCMR-JA (Charles H. Harris-Patent
Atty)
CITY: FORT DETRICK
STATE: MARYLAND
COUNTRY: USA
ZIP: 21702-5012
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,776B
FILING DATE: 01-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/882,431
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Charles H. Harris
REGISTRATION NUMBER: 34,616
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 619-2065
TELEFAX: (301) 619-7714
INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:
LENGTH: 266
TYPE: Amino Acid
STRANDEDNESS: Unknown
TOPOLOGY: Unknown
MOLECULE TYPE: Peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 14:
US-09-144-776B-14
Query Match 21.0%; Score 255.5; DB 4; Length 266;
Best Local Similarity 30.2%; Pred. No. 2.7e-19;
Matches 74; Conservative 45; Mismatches 105; Indels 21; Gaps 9;
QY 1 SEKSEINEKDLRKSELQGTALGNLKOIY--YNSKAITSSEKSAQOFLNTLLFKGFF 58
DB 27 AESQPDPTDELHKASKFTG-LMENMKVLYDDHYVS---ATKVKSVDFRAHDLIYNISD 82
QY 59 TGHWPYNLLVDLGGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTCMYGGV 111
DB 83 KKLKNYDKVKTTELLNEGLAKKYKDEVVDVGYGSNNYVNCYFSSKDNVGVGTGKTCMYGGI 142
QY 112 TLHDNNRLTEE--KKVPINLWIDGKQTPIDKVKTSKKEVTVOELDLQARHYLHGKFG 169
DB 143 TKHEGHNFDGNLQNLVLRVY-ENKRNITISFE-VQTDKKSVTAQOELDIKARNFLINKKNL 200
QY 170 YNSDSFGGKVGORGLIVFHSSEGSTVSYDLFDAQGOYPD--TLRLYRDNTTISSTLSLS 227
DB 201 YEFNS--SFYETGYIKFNIENNGNTFWYDMMPAPGDKFQDSKYLMMYNDNKNVDSKSVKIE 258
QY 228 LLYLT 232
DB 259 VHLTT 263

RESULT 46
US-09-144-776B-8
Sequence 8, Application US/09144776B
Patent No. 6399332
GENERAL INFORMATION:
APPLICANT: Robert G. Ulrich,
Mark A. Olson
Sina Bavari
TITLE OF INVENTION: Bacterial Superantigen
Vaccines
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles H. Harris
STREET: US Army MRC -504 Scott Street
MCMR-JA (Charles H. Harris-Patent
Atty)
CITY: FORT DETRICK
STATE: MARYLAND
COUNTRY: USA
ZIP: 21702-5012
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.5
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144,776B
FILING DATE: 01-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/882,431
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Charles H. Harris
REGISTRATION NUMBER: 34,616
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 619-2065
TELEFAX: (301) 619-7714

```
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 266
; TYPE: Amino Acid
; STRANDEDNESS: Unknown
; TOPOLOGY: Unknown
; MOLECULE TYPE: Peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-144-776B-8
```

```
Query Match 20.6%; Score 250.5; DB 4; Length 266;
Best Local Similarity 30.7%; Pred. No. 9.4e-19;
Matches 75; Conservative 43; Mismatches 107; Indels 19; Gaps 8;

QY 1 SEKSEINEKDLRKKSELOQTALGNLKOIYYNKSATTSSEKSADQFLNTLLFKGFFTG 60
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
27 AESQDPKPDELHKSSKFTG-LMENNK-VLYDDNHVSAINVKSIDQFRYFDLIYSIKDTK 84
QY 61 HPWYNDLLDLGSTAATSEYEGSSVDLYGAYGYQCAG-----TPNKTACMYGG 110
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
85 LGYDNRVVEFKNKDLADKYDKYVDVEGANAYYQCAFSSKKTNDINSHQTDKRTKCMYGG 144
QY 111 VTLDHNNRLTEKKVPINLWIDGKQTVPIDKVKTSKEVTVQELDLQARHYLHGKFG 170
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
145 VTEHGNQLDKYRSITVRVFDGK-NLLSFD-VQYNNKKVTAQELDYLTRHYLVKNKKLY 202
QY 171 NSDSFGGKVRGLIVFHSSEGSTVSYDLFDAQGOYPD--TLIRIYRDNNTTISSTLS 228
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
203 ERNN--SPYETGYIKFIENENS-FWYDMPAPGDKFDQSKYLMYNNDKWDSKDVKIEV 259
QY 229 VLYT 232
Db : : : : :
260 VLYT 263
```

```
RESULT 47
US-08-896-933-21
; Sequence 21, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005002
; CURRENT APPLICATION NUMBER: US/08/896.933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-08-896-933-21
```

```
Query Match 18.5%; Score 225.5; DB 4; Length 239;
Best Local Similarity 30.2%; Pred. No. 3.8e-16;
Matches 74; Conservative 43; Mismatches 105; Indels 23; Gaps 11;

QY 2 EKSEINEKDLRKKSELOQTALGNLKOIYYNKSATTSSEKSADQFLNTLLFKGFFTG 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 ESQDPKPDELHKSSKFTG-LMENNK-VLYNNDHVSAINVKSINEFFDLIYSIKDTKL 58
QY 62 PWYNDLLDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGG 110
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
59 GNYDNRVVEFKNKDLADKYDKYVDVFGANY-YOCYFSKKTNNIDSHENTKRKT-CMYG 116
QY 111 VTLDHNNRLTE-EKKVPINLWIDGKQTVPIDKVKTSKEVTVQELDLQARHYLHGK 169
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 VTEHGNQLDKYRSITVRVFDGK-NLLSFD-VQYNNKKVTAQELDYLTRHYLVKNKK 174
QY 170 YNSDSFGGKVRGLIVFHSSEGSTVSYDLFDAQGOYPD--TLIRIYRDNNTTISST 227
```

```
Db 175 YEFNN--SPYETGYIKFIENENS-FWYDMPAPGKNKFDQSKYLMYNNDKWDSKDV 231
QY 228 VLYT 232
Db 232 VLYT 236

RESULT 48
US-09-314-235-21
; Sequence 21, Application US/09314235
; Patent No. 6338845
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005004
; CURRENT APPLICATION NUMBER: US/09/314,235
; EARLIER FILING DATE: 1999-05-18
; EARLIER APPLICATION NUMBER: 08/896,933
; EARLIER FILING DATE: 1997-07-18
; EARLIER APPLICATION NUMBER: 08/252,978
; EARLIER FILING DATE: 1994-06-02
; EARLIER APPLICATION NUMBER: 07/891,718
; EARLIER FILING DATE: 1992-06-01
; EARLIER APPLICATION NUMBER: US91/00342
; EARLIER FILING DATE: 1991-01-17
; EARLIER APPLICATION NUMBER: 07/466,577
; EARLIER APPLICATION NUMBER: 1990-01-17
; EARLIER APPLICATION NUMBER: 07/416,530
; EARLIER FILING DATE: 1989-10-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-314-235-21
```

```
Query Match 18.5%; Score 225.5; DB 4; Length 239;
Best Local Similarity 30.2%; Pred. No. 3.8e-16;
Matches 74; Conservative 43; Mismatches 105; Indels 23; Gaps 11;

QY 2 EKSEINEKDLRKKSELOQTALGNLKOIYYNKSATTSSEKSADQFLNTLLFKGFFTG 61
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 ESQDPKPDELHKSSKFTG-LMENNK-VLYNNDHVSAINVKSINEFFDLIYSIKDTKL 58
QY 62 PWYNDLLDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGG 110
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
59 GNYDNRVVEFKNKDLADKYDKYVDVFGANY-YOCYFSKKTNNIDSHENTKRKT-CMYG 116
QY 111 VTLDHNNRLTE-EKKVPINLWIDGKQTVPIDKVKTSKEVTVQELDLQARHYLHGK 169
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
117 VTEHGNQLDKYRSITVRVFDGK-NLLSFD-VQYNNKKVTAQELDYLTRHYLVKNKK 174
QY 170 YNSDSFGGKVRGLIVFHSSEGSTVSYDLFDAQGOYPD--TLIRIYRDNNTTISST 227
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
175 YEFNN--SPYETGYIKFIENENS-FWYDMPAPGKNKFDQSKYLMYNNDKWDSKDV 231
QY 228 VLYT 232
Db 232 VLYT 236

RESULT 49
US-08-896-933-20
; Sequence 20, Application US/08896933
; Patent No. 6221351
; GENERAL INFORMATION:
; APPLICANT: Terman, David S.
; TITLE OF INVENTION: TUMOR KILLING EFFECTS OF ENTEROTOXINS,
; TITLE OF INVENTION: SUPERANTIGENS, AND RELATED COMPOUNDS
; FILE REFERENCE: 09629/005002
```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:58:32 ; Search time 4.3768 Seconds
(without alignments)
1032.821 Million cell updates/sec

Title: US-09-900-766-2

Perfect score: 1218

Sequence: 1 SEKSEINEKDLRKKELOG.....RDNTTISLSISLYLTT 233

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 118974 seqs, 19401057 residues

Total number of hits satisfying chosen parameters: 118974

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1072	88.0	248	9	US-09-870-759-16
2	935	76.8	257	9	US-09-870-759-8
3	611	50.2	258	9	US-09-870-759-14
4	268.5	22.0	251	8	US-08-973-391A-13
5	268.5	22.0	266	9	US-09-870-759-10
6	267.5	22.0	239	10	US-09-150-947B-12
7	263.5	21.6	266	9	US-09-870-759-12
8	220	18.1	250	9	US-09-870-759-20
9	100	8.2	234	9	US-09-870-759-18
10	87	7.1	226	10	US-09-815-242-5900
11	87	7.1	226	10	US-09-815-242-13156
12	82	6.7	772	9	US-10-121-032-28
13	79	6.5	833	10	US-09-844-281-1
14	78	6.4	607	10	US-09-815-242-13379
15	78	6.4	607	10	US-09-815-242-13682
16	78	6.4	1551	10	US-09-864-761-35904
17	77.5	6.4	312	10	US-09-815-242-4904
18	77.5	6.4	312	10	US-09-815-242-10541
19	77.5	6.4	985	9	US-09-738-626-4377
20	77.5	6.4	985	9	US-09-738-626-4377

20	77.5	6.4	5795	10	US-09-815-242-12610	Sequence 12610, A
21	77	6.3	579	9	US-10-108-605-215	Sequence 215, App
22	77	6.3	704	10	US-09-801-368-218	Sequence 218, App
23	76.5	6.3	454	10	US-09-815-242-13445	Sequence 13445, A
24	75.5	6.2	424	9	US-09-738-626-4306	Sequence 4306, Ap
25	75.5	6.2	818	9	US-10-055-364-44	Sequence 44, Appl
26	74.5	6.1	386	9	US-09-738-626-4498	Sequence 4498, Ap
27	74.5	6.1	667	10	US-09-815-242-10709	Sequence 10709, A
28	74	6.1	750	10	US-09-815-242-13405	Sequence 13405, A
29	73.5	6.0	882	10	US-09-815-242-10668	Sequence 10668, A
30	73.5	6.0	1158	10	US-09-815-242-4994	Sequence 4994, Ap
31	73.5	6.0	1207	10	US-09-815-242-10626	Sequence 10626, A
32	73	6.0	275	10	US-09-054-141-2	Sequence 2, Appli
33	73	6.0	320	10	US-09-740-452-2	Sequence 2, Appli
34	73	6.0	320	10	US-09-815-242-13278	Sequence 13278, A
35	73	6.0	320	10	US-09-815-242-13621	Sequence 13621, A
36	73	6.0	343	10	US-09-939-521-8	Sequence 8, Appli
37	73	6.0	357	10	US-09-925-301-970	Sequence 970, App
38	73	6.0	361	12	US-10-109-885-8	Sequence 8, Appli
39	73	6.0	510	10	US-09-205-658-45	Sequence 45, Appl
40	73	6.0	510	10	US-09-844-353A-45	Sequence 45, Appl
41	73	6.0	566	10	US-09-801-368-418	Sequence 418, App
42	73	6.0	893	9	US-10-014-436-4	Sequence 4, Appli
43	73	6.0	934	8	US-08-837-459-19	Sequence 19, Appl
44	73	6.0	1155	10	US-09-801-368-304	Sequence 304, App
45	73	6.0	1737	9	US-09-808-602-83	Sequence 83, Appl

ALIGNMENTS

RESULT 1
US-09-870-759-16
; Sequence 16, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 248
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-870-759-16

Query Match 88.0%; Score 1072; DB 9; Length 248;
Best Local Similarity 90.2%; Pred. No. 2.1e-100;
Matches 202; Conservative 9; Mismatches 13; Indels 0; Gaps 0;

Qy	1	SEKSEINEKDLRKKELOGTALGNLQIYYNSKAITSEKSADQFLNTLLFKGFTG	60
Db	25	SEKSEINEKDLRKKELOGTALGNLQIYYNEKAITENKESDDQFLNTLLFKGFTG	84
Qy	61	HPWYNLLVDLGLSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTCMYGGVTLHDNNRLT	120
Db	85	HPWYNLLVDLGLSKDATNKYKGGKVDLYGAYGYOCAGGTPNKTCMYGGVTLHDNNRLT	144
Qy	121	EKKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHLYHGKFGLYNSDSFGCKVQ	180
Db	145	EKKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHLYHGKFGLYNSDSFGCKVQ	204
Qy	181	RGLIVFHSSEGSTVSYDLFDAQGYPDTLRLRIYRDNNTTISSTSL	224
Db	205	RGLIVFHSSEGSTVSYDLFDAQGYPDTLRLRIYRDNNTTISSTSL	248

RESULT 2

```
US-09-870-759-8
; Sequence 8, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 257
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
US-09-870-759-8

Query Match 76.8%; Score 935; DB 9; Length 257;
Best Local Similarity 75.5%; Pred. No. 1.3e-86;
Matches 176; Conservative 21; Mismatches 36; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKKSELQGTALGNLKOIYYNNSKAITSEKSAOQFLNTLLFKGFFTG 60
Db 25 SEKSEINEKDLRKKSELQGTALGNLKOIYYNNEKAKTENKESHQDLQHTILFKGFFTN 84
QY 61 HPWYNDLLVDLSTAASTSEYESSVDLYGAYGYQCAGGTPNKTCACMYGGVTLHDNRLLT 120
Db 85 HSWYNDLLVDFDSKIDVYKKGKVDLYGAYGYQCAGGTPNKTCACMYGGVTLHDNRLLT 144
QY 121 EEKKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFGGKVQ 180
Db 145 EEKKVPINLWIDGKQTTVPLETKNNKNTVQELDLQARRYLQEKYNYNSDSVDFGKVQ 204
QY 181 RGLIVPHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNTTISSTSLSLYLTT 233
Db 205 RGLIVPTSTEPSVNYDLFGAQQNSNTLLRIYRDNTTINSENMHIDILYLT 257

RESULT 3
US-09-870-759-14
; Sequence 14, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 258
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
US-09-870-759-14

Query Match 50.2%; Score 611; DB 9; Length 258;
Best Local Similarity 51.1%; Pred. No. 5.3e-54;
Matches 118; Conservative 35; Mismatches 78; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKKSELQGTALGNLKOIYYNNSKAITSEKSAOQFLNTLLFKGFFTG 60
Db 26 NENIDSVKEKELHKKSELSTALNNMKHSYADKNPIIGENKSTGDFLENTLLYKFFTD 85
QY 61 HPWYNDLLVDLSTAASTSEYESSVDLYGAYGYQCAGGTPNKTCACMYGGVTLHDNRLLT 120
Db 86 LNFEDLLINFNSKEMAQHFKNVDVYPIRSYNGYGEIDRTACTYGGVTPHEGNK 145
QY 121 EEKKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFGGKVQ 180

US-09-870-759-10
; Sequence 10, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 266
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
US-09-870-759-10

Query Match 22.0%; Score 268.5; DB 8; Length 251;
Best Local Similarity 32.5%; Pred. No. 1.4e-19;
Matches 77; Conservative 45; Mismatches 96; Indels 19; Gaps 10;

QY 4 SEEINEKDLRKKSELQGTAL-GNLKOIYY-YNSKAIT-SSSEKSAOQFLNTLLFKGFFTG 60
Db 25 SQEVFAQQDPDFSQLHRSSLVKNLQNIYFLYEGDPVTHENKSVQDOLLSHHLYN---VS 81
QY 61 HPWYNDLLVDLSTAASTSEYESSVDLYGAYGYQC-AGGTPNKTCACMYGGVTLHDNRLL 119
Db 82 GPNYDKLKTENKQEMATLFKDKNDIYGVYVYHLCYLCENASACIYGGVTNHEGNHL 141
QY 120 TEKKVPINLWIDGKQTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFGGKV 179
Db 142 EIPKXIVKVSIDGIQ-SLSFD-IETNKKMVTQAQELDYKRVKYLTDNKKOLYTNGP--SKY 197
QY 180 QRGLIVPHSSEGSTVSYDLFD---AQGYPTDLLRIYRDNTTISSTSLSLYLTT 232
Db 198 ETGYIKFIPKNKESFWDFPEPEFTQSKY----LMIYKDNETLDSNTSQIEVILTT 250

RESULT 5
US-09-870-759-10
; Sequence 10, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 266
; TYPE: PRP
; ORGANISM: Staphylococcus aureus
US-09-870-759-10

Db 146 ERKKIPINLWINGVQKEVSLDKVQTDKKNVTYQELDAQARRYLQDKLKYNDTLGGKIQ 205
QY 181 RGLIVPHSSEGSTVSYDLFDAQOQYPTDLLRIYRDNTTISSTSLSLYLTT 231
Db 206 RGKIEFDSSDGSKSVSYDLFDVKGDFPEKQLRIYSDNKTLSLSTELHLDIYLY 256
```



```
Db 35 KPSQLORSNLVTKFIYIFMRVTLVTHENVKSVSDQLSHDLIYN---VSGPNYDKLKTE 91
QY 71 LGSTAATSEYEGSSVDLYGAYGYQC-AGTGNKTACMYGGVTLHDNNRLTEKKVPINL 129
Db 92 LKQEMATLFDKKNVDIYVEVYHCLYCENASACLYGGVTNHGHNLEIPKRIIVKV 151
QY 130 WIDKGT-TVPIDKVKTSKKEVTQVQLDLQARHYLHGKFGLYNSDSFGKGQGLIVFHS 188
Db 152 SIDGIOSLFDIEQKNG-----NCSRISTVVRKYLTDNKLQLTNGP--SKYETGIYKIFP 205
QY 189 SEGTSVSYDLFD-----AQOQYPTDLRIYRDNTTISSTLSISLYLT 232
Db 206 KKNESWFDFEPPEPTQSKY-----LMIVKNETLDSNTSQIEVLYLT 249

RESULT 9
US-09-870-759-18
; Sequence 18, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 18
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-870-759-18

Query Match 8.2%; Score 100; DB 9; Length 234;
Best Local Similarity 24.5%; Pred. No. 0.012;
Matches 48; Conservative 33; Mismatches 61; Indels 54; Gaps 12;

QY 25 NLKQI---YYNSKATTSSE-----KSADQFLNTLLFKGFFTGHPWYNDLL-- 58
Db 45 NIKDLLDWSYSSGSDTFTNSEVLNDSLGSMRIKNTDGSI-SLIIFPS-----PYISPAFTK 98
QY 69 ---VDLGGTAATSEYEGSSVDLYGAYGYQCAGTGNKTACMYGGVTLHDNNRLTEKKV 125
Db 99 GEKVDL-NTRKTKSHTS---ECTIHFQISGVT-----NTEKLPTPIEL 140
QY 126 PINLWIDGKQTTVPIDKVKTSKKEVTQVQLDLQARHYLHGKFGLY-NSDSFGG--KVQRG 182
Db 141 PLKVKVHGKDSPLKYGP-RFDKQLAISTLDPEIRHQLTQIHGLYRSSDKTGGYWKIT-- 197
QY 183 LIVFHSSEGSTVSVDL 198
Db 198 -----MNDGSTVQSGL 208

RESULT 10
US-09-815-242-5900
; Sequence 5900, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; CURRENT APPLICATION NUMBER: US/09/815,242
```

```
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5900
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5900

Query Match 7.1%; Score 87; DB 10; Length 236;
Best Local Similarity 21.9%; Pred. No. 0.23;
Matches 42; Conservative 42; Mismatches 92; Indels 16; Gaps 6;

QY 8 NEKDRLKSELQGTALGNLKIYYNSKAITSEKSADQFLNTLLFKGFFTGHPWYNDL 67
Db 24 NVQSVQAKAEVQOSESELK--HYTNKPILEKRNVTGFKYTDGKHYLEVTVGQ--QHSR 79
QY 68 LVDLGSTA-ATSEYEGSSVDLYGAYGYQCAGTGNKTACMYGGVTLHDNNRLTEKKVP 126
Db 80 ITLLGSDKDKFKDGENSNIDVF-----ILREGDSQOATWYSGVTKSVSYQIDINTP 134
QY 127 INLWIDGKQTTVPIDKVKTSKKEVTQVQLDLQARHYLHGKFGLYNSDSFGKGQGLIVF 186
Db 135 I-LEIKKQNEVDLKFYIYSKEDISLKELDYRLRERAIKQHGLYSNG-----LKQGOITI 188
QY 187 HSSEGSTVSVDL 198
Db 189 TWDGTTHTIDL 200

RESULT 11
US-09-815-242-13156
; Sequence 13156, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; CURRENT APPLICATION NUMBER: 60/257,931
```



```
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13156
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-13156

Query Match          7.1%; Score 87; DB 10; Length 226;
Best Local Similarity 21.9%; Pred. No. 0.23;
Matches 42; Conservative 42; Mismatches 92; Indels 16; Gaps 6;

QY 8 NEKDLRKLSELOCTALGNLKOIYYNSKAITSSEKSADQFLNTLLFKGFFGHPWYNDL 67
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 24 NVQSVQAKAEVQKQSESELK--HYNKPILERKNVTGKYTDEGRHYLEVTVGQ--QHSR 79
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 68 LVDLGSTA-ATSEYEGSSVDLYGAYGYQCAGGTPNKTACHYGGVTLHDNNRLTEKKVP 126
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 80 ITLLGSDKDKFKDGENSIDVE----ILREGDSRQATNYSIGGVTKSNSVQYIDYINTP 134

QY 127 INLWIDGKQTTVPIDKVTSEKVTVOELDLQARHVLHGKFGLYNSDSFGGKVQRGLIVF 186
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 135 I-LEIKKONEDVLKDFYISKEDISKELDYRLRERAIKQHGLYSNG-----LKQGQITI 188
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 187 HSSEGSTSVSYDL 198
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 189 TWNQGTHTIDL 200

RESULT 12
US-10-121-032-28
; Sequence 28, Application US/10121032
; Patent No. US2002015550A1
; GENERAL INFORMATION:
; APPLICANT: Bylina, Edward J.
; TITLE OF INVENTION: GLYCOSIDASE ENZYMES
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Gray Cary Ware & Freidenrich LLP
; STREET: 4365 Executive Drive, Suite 1600
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92121
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/121,032
; FILING DATE: 09-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,078
; FILING DATE: 13-AUG-1998
; APPLICATION NUMBER: 08/949,026
; FILING DATE: 10-OCT-1997
; APPLICATION NUMBER: 60/056,916
; FILING DATE: 06-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/024002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 858/677-1456
; TELEFAX: 858/677-1465
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 772 amino acids
```

```
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 28:
US-10-121-032-28

Query Match          6.7%; Score 82; DB 9; Length 772;
Best Local Similarity 28.3%; Pred. No. 3.9;
Matches 32; Conservative 15; Mismatches 36; Indels 30; Gaps 6;

QY 4 SEINEKDLRKXSEL--QG-----TALGNLKOIYYNSK--AITSEKSADQFLNTLLP 54
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 115 SESLEEDLRKVELTIEGYPARVIMBEILDYDYGELGAVYSPEKT-----IF 165
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 55 KGFFTGHPWYNDLLVDLGS----TAATSEYEGSSV-----DLYGAYGYQ 95
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 166 RVWSPYSKWKVKLLFKNGEDTEPYQVNNMEYKNGVWEAVVEGDLGDFVLIYQ 218

RESULT 13
US-09-844-281-1
; Sequence 1, Application US/09844281
; Patent No. US20020082386A1
; GENERAL INFORMATION:
; APPLICANT: Mangold, Beverly L.
; APPLICANT: Aldrich, Jennifer L.
; APPLICANT: O'Brien, Thomas
; TITLE OF INVENTION: Anthrax Specific Antibodies
; FILE REFERENCE: 38602.0003
; CURRENT APPLICATION NUMBER: US/09/844,281
; CURRENT FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/200,505
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Patentcin version 3.1
; SEQ ID NO 1
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Bacillus anthracis
US-09-844-281-1

Query Match          6.5%; Score 79; DB 10; Length 833;
Best Local Similarity 22.6%; Pred. No. 8.8;
Matches 55; Conservative 29; Mismatches 97; Indels 62; Gaps 10;

QY 33 NSKATISSEKSADQFLNTLLFKGFFTGHPWYNDLLVDLGSTAATSEYEGSSVDLYGAY 92
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 403 NSONLVVGEKASLNKLVIATIGE-----DKYVDPGSIKSSNHG-IISVVNNYI 451

QY 93 GYQCAGGTPNKTACMYGGVT-----LHDNNRLTEKKVPINLWIDGKQTTVPIDKVK 145
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 452 TAEAGEA--TLTIKVGDTVTKDKFKVTTDSRKLVSVKANPKLQV--VONKTLPTVFTVT 508

QY 146 SK-----REV-----TVQELDLQARHYLHGKFGLYNSDSFGGKVQGRGLIVFHS 188
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 509 DOYGDPPFGATAAIKEVLPKTVGVAEGGLDVTTDSGSIKTKTIGVTGNDVGEQVHFQN 568

QY 189 SEGSTV-----SYDLFDAQGY---PDTLLRIYRDNNTTISSTLSLSLYL 230
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 569 GNGATLGLSLVNVNTEGNVAFKFNELSVKGYGQSQSPDTKLDLN-----VSTTVEYQLSK 622

QY 231 YTT 233
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 623 YTS 625

RESULT 14
US-09-815-242-13379
; Sequence 13379, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
```



```
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 35904
; LENGTH: 1551
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL050341.18
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1
; OTHER INFORMATION: EST_HUMAN HIT: AV715688.1, EVALUE 1.00e-129
; OTHER INFORMATION: SWISSPROT HIT: Q06730, EVALUE 5.00e-15
US-09-864-761-35904

Query Match 6.4%; Score 78; DB 10; Length 1551;
Best Local Similarity 23.9%; Pred. No. 26;
Matches 51; Conservative 31; Mismatches 55; Indels 76; Gaps 11;

QY 7 INEKDLRKSELQGTALGNLKIYYNS--KAITSEKSADQF-----LNTLTLF 54
DB 1382 VPKENFRKHQSPRDLTKYKPMGFESSFLKFIQSEEEKDFDDWEPSEHITLNSOSS 1441
QY 55 KGFFTGHPWYNLLVDLGSTAAATSEYSGSVLDYGAAYGYQCAGGTPNKTACMYGGVTLH 114
DB 1442 SNDLTGNVANNVND-----SEPE---VDI-----PHSSS-----DSTTH 1474
QY 115 DNNRLTEEKVPINLWDGKQTTVP-----IDKVKTSKKEVTVQVQLDQARHYLHCK 166
DB 1475 EN--LT-----AIPPLIVAETTVFSLNLRVLDKALDCGELAKQL-----HYL---- 1519
QY 167 FGLYNSDSFGKVGKQLIVFHSSEGSTVSYDLF 199
DB 1520 -----RPVVVLERSKESFTPLDLF 1538

RESULT 17
US-09-815-242-4904
; Sequence 4904, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10541
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10541

Query Match 6.4%; Score 77.5; DB 10; Length 312;
Best Local Similarity 34.0%; Pred. No. 3.2;
Matches 16; Conservative 12; Mismatches 16; Indels 3; Gaps 1;

QY 119 LPEKKVPINLWDGKQTTVPIDKVKTSKK---EVTVQVQLDQARHY 162
DB 215 MTEHLDFVNLDEAKNAEIMVKEETQKEKMLEMTEELDLSVRSY 261

RESULT 18
US-09-815-242-10541
; Sequence 10541, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10541
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Enterococcus faecalis
US-09-815-242-10541

Query Match 6.4%; Score 77.5; DB 10; Length 312;
Best Local Similarity 34.0%; Pred. No. 3.2;
```



```

; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; PRIOR FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4306
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4306

Query Match 6.2% Score 75.5; DB 9; Length 424;
Best Local Similarity 20.8%; Pred. No. 7.8;
Matches 38; Conservative 26; Mismatches 56; Indels 53; Gaps 9;

Qy 74 TAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGVTLHDNNRLT---EKKVPIPLNW 130
Db 196 TTQLKQVEGLSVNTIGFEG-----WGGSVLDDGNGVTVDSDAKAGLQAL 241

Qy 131 IDG-KQTVPTDKVTKSKKEVTVOELDLQARH-----YLF-----GKF---GLYN 171
Db 242 VDGDDGTISKASLAATEENLAFTEGQTAYAINWPYMTNSPEAEATAGKFEVQPLVG 301

Qy 172 SDSFG---GRVQGLIVFHSSEGSTVSVDLFD-----AQQYPTDLLRIYRDN 216
Db 302 KDGVGSTLGGYNGINV--NSENKATARDPIEFIEINEQTFWADNSFPVPLASIYDDE 359

Qy 217 TTI 219
Db 360 SLV 362

RESULT 25
US-10-055-364-44
; Sequence 44, Application US/10055364
; Patent No. US20020155433A1
; GENERAL INFORMATION:
; APPLICANT: Patience, Clive
; TITLE OF INVENTION: Gamma Herpes Virus DNA and Methods of Use
; FILE REFERENCE: 61750-379
; CURRENT APPLICATION NUMBER: US/10/055,364
; CURRENT FILING DATE: 2002-01-23
; PRIOR APPLICATION NUMBER: US/09/612,204
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US/60/142,736
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: US/60/168,532
; PRIOR FILING DATE: 1999-12-02
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 44
; LENGTH: 818
; TYPE: PRT
; ORGANISM: Equine herpesvirus 5
; US-10-055-364-44

Query Match 6.2% Score 75.5; DB 9; Length 818;
Best Local Similarity 22.3%; Pred. No. 19;

```

```

Matches 52; Conservative 33; Mismatches 81; Indels 67; Gaps 13;

Qy 6 EINEKDLRKKSELOQTALGNLKOIYYVNSKAITSSSEKSAQDLFTNTLLFKGFFTGHPWYN 65
Db 73 EENCPTNTEDEHVEGILL-----IYKTNIVPYIFNVKRYKRLVSTIYKG-----WSQ 121

Qy 66 DLLVDLGSTAATSEYEGS-SVDLYGAY---YGYQCAGGTPNKTACMYGVTLHDNNRLT 121
Db 122 D-----AITNQYTSFAMPLWEARLVDVNYEC-----YNGIQVTENGHLT- 161

Qy 122 EKKVPIPLNWIDGKQTTVPIDKVKTSKEVTVOELDLQARHLYHCKFGLYNSDSFGGQVOR 181
Db 162 -----TYVDRDGYNESVRLVPADGLTSSIRRY-HSQPELY-----VTPR 199

Qy 182 GLIVFHSSEGSTVSVDLFD--AQQYPTDLLRIYRDNNTTISSTSLISLVLVT 232
Db 200 NLLWSYTR-TTVNCEVIDMTARSHKP-----FEYFVTASGDSIETSPF-YT 244

RESULT 26
US-09-738-626-4498
; Sequence 4498, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIALI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENO, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4498
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
; US-09-738-626-4498

Query Match 6.1% Score 74.5; DB 9; Length 386;
Best Local Similarity 21.4%; Pred. No. 8.6;
Matches 41; Conservative 23; Mismatches 67; Indels 61; Gaps 8;

Qy 23 LGNLKQIYYVNSKA-ITSSEKSAQDLFTNTLLFKGFFTGHP-----WYNLLDVD 70
Db 131 LVNSALVYYSSNAFLHAAEKLGTSSFSHTFVSGEVTGIPYELEDCGFTFYERDLAE 190

Qy 71 LG-----STAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMY-----GGVT-- 112
Db 191 LGPHFDRLPEDMKAFALNDGGDRLEQFEQINSFVC---RPRHTGNQFLAMQTKGAKTPY 247

Qy 113 -----LHDNNRLTEKKVPIPLNWIDGKQTTVPIDKVKTSKEVTVOELDLQARHLYH 165
Db 248 IPRHFHRLHTENFIC-----LDGR-----VKLHVNGQEITLSRGDVYHA 286

Qy 166 KFGLYNSDSFGG 177
Db 287 PAGTIHSFAFAG 298

```

RESULT 27

US-09-815-242-10709
; Sequence 10709, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 10709

; LENGTH: 667

; TYPE: PRT

; ORGANISM: Enterococcus faecalis

US-09-815-242-10709

Query Match 6.1%; Score 74.5; DB 10; Length 667;

Best Local Similarity 17.8%; Pred. No. 18;

Matches 63; Conservative 36; Mismatches 86; Indels 169; Gaps 14;

QY 16 SELOGTALCN-----LKOIYYNSKAITSSSEKSADQFLTNL-----LTKGTF 58

DB 263 SSVHGAPICEEITAAKAVGHEYPDFTVPVEVAAFRKFTMIDEGQKAEANMEFKNYE 322

QY 59 TGHF-----WYNDL-LVDLGSTAA----- 76

DB 323 HARPELAKQKFAFANQLPEGEQELPKYELGTSAASRVTSKETIAISKVVPSEWGGSA 382

QY 77 -----TSEYEGSSV-----DLYGAYYGQCAGGTPNKTACHMYGGV- 111

DB 383 DLSASNTMVAEKFQESQYEGRWLWFGVREFAFAMAAMNGTQLHGSSH-----VYGGTF 437

QY 112 -----TLHDNNRLTEKKVPI-----NLWT----- 131

DB 438 FVFTDYLRAIRLAALQKVPVTVLTHDSVAVGEDGTPHTEIQLASVRCIPNVHVIRPA 497

QY 132 DGKQTTVPIDKVKTSKKEVTQVQELQARHYLHGKFGLYNSDSFGKVGQGLIVFHSSEG 191

DB 498 DGNETVAANKIAMTSTETPTILVLSRQNLPLVLEGTLE-HASDS-----VQKAYVLSPQKG 552

QY 192 -----STVSVDLFDQAQ-GQYPTDLL 210

DB 553 EQPSGILITATSGEVNLAVEAQAALAEEDIDVSVWSPSPDLFEKQSAEYKESVL 606

RESULT 28

US-09-815-242-13405
; Sequence 13405, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

; TITLE OF INVENTION: Identification of Essential Genes in

; FILE REFERENCE: ELITRA.011A

; CURRENT APPLICATION NUMBER: US/09/815,242

; CURRENT FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: 60/191,078

; PRIOR FILING DATE: 2000-03-21

; PRIOR APPLICATION NUMBER: 60/206,848

; PRIOR FILING DATE: 2000-05-23

; PRIOR APPLICATION NUMBER: 60/207,727

; PRIOR FILING DATE: 2000-05-26

; PRIOR APPLICATION NUMBER: 60/242,578

; PRIOR FILING DATE: 2000-10-23

; PRIOR APPLICATION NUMBER: 60/253,625

; PRIOR FILING DATE: 2000-11-27

; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22

; PRIOR APPLICATION NUMBER: 60/269,308

; PRIOR FILING DATE: 2001-02-16

; NUMBER OF SEQ ID NOS: 14110

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 13405

; LENGTH: 750

; TYPE: PRT

; ORGANISM: Streptococcus pneumoniae

US-09-815-242-13405

Query Match 6.1%; Score 74; DB 10; Length 750;

Best Local Similarity 22.0%; Pred. No. 24;

Matches 53; Conservative 34; Mismatches 90; Indels 64; Gaps 11;

QY 26 LKOIYYNSKAITSSSEKSADQFLTN-----TLTKGFFTGHPHY-----NDLLVDLGST 74

DB 328 LAQIVDFVKKAQTS--RAPIDBLTKISGIFVPVVVILGIMTFWVFLRLDSVVVLGAS 385

QY 75 AATSEYEGSSVDLYGAYYGQCAGGTPNKTACHMYG-----GVTLDNNRLTEKKVPI 127

DB 386 FVSSLLYGAVALLIAC----PCALGLATPTALMGVGRSAKMGVLLKNGTVLQETOKVQT 441

QY 128 NLW-----IDGKQTTVPI--DKVRTSKKEVTQVE-----LDLQA-- 159

DB 442 LVFDKGTGTLECKPVPVTDIIGDEVEVFGLAASLEDAHQPLAEAIKVRASEAGLEFQTV 501

QY 160 -RHYLHGKFGCLYNSDSFGKVGQGLIVFHSSEGSTVSYDLFDAQGOYPTDLLRIYRDNTT 218

DB 502 NFOALHGK-----GVSGRINGQVLLH-----GNAKMLDGMDSINTYQDKLEELEAKT 550

QY 219 I 219

DB 551 V 551

RESULT 29

US-09-815-242-10668
; Sequence 10668, Application US/09815242
; Patent No. US20020061569A1

GENERAL INFORMATION:

; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 10668
LENGTH: 882
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10668

Query Match 6.0%; Score 73.5; DB 10; Length 882;
Best Local Similarity 19.1%; Pred. No. 34;
Matches 50; Conservative 41; Mismatches 112; Indels 59; Gaps 9;
Qy 16 SELQGTALGNLKOIYYNSKAITSSEKSADQFLNTL-----LFGGFTG 60
Db 335 SPIVGLANGNNKMYTNNLAVFESQFIDLWMDTRKKNVYDAKRYVALNRVVGKMTG 394
Qy 61 HPW-----YNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTGPNKTACHYGG 110
Db 395 IAFDVLLAAVLLDTNDNADIEGVAHQYDAIQSD--EAIYKGAKKGLPEDEEVFFGH 452
Qy 111 -----VTLHDNNRLTEEK-----KVPINLWI-DGKOTTPIDKVKTSKKEVT 151
Db 453 LARKIKAIQFLTSLKDLSELTEKNQADLYFKMELPLSRILGDMETITGIRVDATRLKEMQVE 512
Qy 152 VQELDQARHYLHGKFGLYNSDSFGGKVQVGRGLIVFH-----SSEGSVSVYDLFDQAQ 202
Db 513 FSERLKEIEKIYAEG--EFNLSNPKQLGVLIFKMGPLVIAKKTKGTSTAVDVLEQL 570
Qy 203 GQYPTDLLRI--YRDNTTISST 222
Db 571 KEQAPIVADILTYRQIAKIQT 592

RESULT 30
US-09-815-242-4994
Sequence 4994, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 4994
LENGTH: 1158
TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-4994

Query Match 6.0%; Score 73.5; DB 10; Length 1158;
Best Local Similarity 24.4%; Pred. No. 49;
Matches 31; Conservative 16; Mismatches 65; Indels 15; Gaps 2;
Qy 111 VTLHDNNRLTEEKVPINLWIDGKOTTPIDKVKTSKKEVTVQELDQARHYLHGKFGLY 170
Db 94 VTLRLTNRETGEIKSQEVFFGDFPLMTMGTFIINGAERVIVSQLVRSPGVYFHGKVDKN 153
Qy 171 NSDSFGGKV--ORGLIVPHSSEGSTVSYDLFDQAQOYP-----DTLLRIYRD 215
Db 154 GKEGFGSTVPINRGAWLEMETDAKDIDSYVRIDRTKRKIPLTLVLRALGFGSDDTIFEIFGD 213
Qy 216 NTTISST 222
Db 214 SESLRNT 220

RESULT 31
US-09-815-242-10626
Sequence 10626, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 10626
LENGTH: 1207

TYPE: PRT
ORGANISM: Enterococcus faecalis
US-09-815-242-10626

Query Match 6.0%; Score 73.5; DB 10; Length 1207;
Best Local Similarity 24.4%; Pred. No. 52;
Matches 31; Conservative 16; Mismatches 65; Indels 15; Gaps 2;
QY 111 VTLHNNRLTEKKVPINLWIDGKQTPIDKVKTSKKEVTVOELDLQARHYLHGKFGLY 170
Db 97 VTLRLTNRETGIKQEVFFGDFPLMTENGTFIINGAERVIVSQLVRSPGVFHKVDKN 156
QY 171 NSDSFGKV-ORGLIVFHSSEGSTVSVDLFDAGQGY- - - - -DTLRIYRD 215
Db 157 GREGFSTVIPNRGAWLEMETDAKDISYVRIDTRKIPLTVLVRLGFGSDDTIFEIGD 216
QY 216 NNTIST 222
Db 217 SESLNT 223

RESULT 32
US-09-054-141-2
Sequence 2, Application US/09054141
Patent No. US20010018193A1
GENERAL INFORMATION:
APPLICANT: DOUGAN, Gordon
APPLICANT: FRANKEL, Gad
TITLE OF INVENTION: Antibodies to Intimin-like
TITLE OF INVENTION: Proteins of E. coli
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: HALE and DORR LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/054,141
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/409,452
FILING DATE: 23-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: BAKER, Hollie L.
REGISTRATION NUMBER: 31,321
REFERENCE/DOCKET NUMBER: 102286.319
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 526-6000
TELEFAX: (617) 526-5000
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 275 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-054-141-2

Query Match 6.0%; Score 73; DB 10; Length 275;
Best Local Similarity 21.4%; Pred. No. 7.6;
Matches 47; Conservative 32; Mismatches 77; Indels 64; Gaps 11;
QY 21 TALGNLKIYYNKAITSSEKSAOFTLNTLLFKGFTTGHWPYNDLLVDLGLSTAATSY 80
Db 10 TAVANGKAIKYTVKVMKNGQPVNQSFTSTNF-GMFNGK-----SQTQATTGN 58

QY 81 EG-SSVDLYGAYYGQACGTPNKACMYG-GVTLHNNRLTEKKVPI-----N 128
Db 59 DGRATITLTSSAGKATVSATVSDGAEVKATEVTFDEDKI--DNKVKIIGNVRGELPN 116
QY 129 LMTDGKQTTVPIDKVKTS-----KKEVTVOELDLQARHYLHGKFGLYNSDSFGKV 179
Db 117 IWLQYGF-----KLKASGGDGYTSWYSENTSIATVDASGKVTLNCK----- 158
QY 180 ORGLIVFHSSEG--SVSYDL-----FDAQGYPDTL 209
Db 159 --GSVVIKATSGDKQTVSYTIKAPSYMIVKIVKQAYYADAM 196
RESULT 33
US-09-740-452-2
Sequence 2, Application US/09740452
Patent No. US20010010912A1
GENERAL INFORMATION:
APPLICANT: Wilding, Edwina Imogen
APPLICANT: Black, Michael T.
APPLICANT: Shilling, Lisa K.
APPLICANT: Kosmatka, Anna L.
APPLICANT: Jaworski, Deborah D.
TITLE OF INVENTION: nrdF
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dechert, Price & Rhoads
STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103-2793
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: Fast-SEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/740,452
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/092,437
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Falk, Stephen T
REGISTRATION NUMBER: 36,795
REFERENCE/DOCKET NUMBER: GM10155
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-994-2488
TELEFAX: 215-994-2222
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-740-452-2

Query Match 6.0%; Score 73; DB 10; Length 320;
Best Local Similarity 25.2%; Pred. No. 9.4;
Matches 31; Conservative 18; Mismatches 40; Indels 34; Gaps 6;
QY 12 LRKKSLOCTALGNLKOIYYNKAITSSEKSAOFTLNTLLFKGFTTGHWPY---NDLL 68
Db 130 LOKKAEI-----VNEIYLNGLSPL---EKKVASVLETFELFYSGFPT--PLYLGNKL 177
QY 69 VDLGTAATSEYGGSVOLYGAYYGQACGTPNKACMYGTVTLHNNRLTEKKVPI 128
Db 178 ANVAEIKLIIRDES---VHGTYIGYKFGOLF-----NLPPEEQEKLK 218

```
Qy 129 LWI 131
   |
Db 219 EWM 221

RESULT 34
US-09-815-242-13278
; Sequence 13278, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13278
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13278

Query Match 6.0%; Score 73; DB 10; Length 320;
Best Local Similarity 25.2%; Pred. No. 9.4;
Matches 31; Conservative 18; Mismatches 40; Indels 34; Gaps 6;

Qy 12 LKKSELOGTALGNLKOIYYNNSKAITSEKSADOFNTLLFKGFFTGHPWY---NDLL 68
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 130 LQKRAEI-----VNEIYLGNSPL---EKKVASVLETFLEYSGFFT--PLYYLGNKKL 177
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 69 VDLGSTAATSEYSSVDLYGAYGYOCAGTTPNKTACMYGGVTLHDNNRLETEKKVPIN 128
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 178 ANVAEIIKLIIRDES---VHGTYIGYKFLQGF-----NELPEEEQEKLK 218
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Qy 129 LWI 131
   |
Db 219 EWM 221

RESULT 36
US-09-939-521-8
; Sequence 8, Application US/09939521
; Patent No. US20020119466A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; APPLICANT: Williams, Mark
; APPLICANT: Rudolph-Owen, Laura A.
; TITLE OF INVENTION: 46863, A NOVEL HUMAN METHYLTRANSFERASE
; FILE REFERENCE: MNI-183
; CURRENT APPLICATION NUMBER: US/09/939,521
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/227,867
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 343
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-939-521-8

Query Match 6.0%; Score 73; DB 10; Length 343;
Best Local Similarity 20.9%; Pred. No. 10;
Matches 49; Conservative 39; Mismatches 80; Indels 66; Gaps 10;
```



```
Qy 124 KVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHLYHGKFGLYNSDSFGGKVQRGL 183
Db 240 SNTIE-----TTTQAQLEKSRGAKK-RIKERALMGS--LHSTL---NGNSIAGSIQ--- 285

Qy 184 IVFHSSEGSTVSYDLFD---AQGQY---PDTLLRIRYDNTTISSTLSIS 227
Db 286 -----TISHDLYDDDSMQGAFDNVPSSFRPRTQSNLSPGSSSRVS 326

RESULT 40
US-09-844-353A-45
; Sequence 45, Application US/09844353A
; Patent No. US20020037585A1
; GENERAL INFORMATION:
; APPLICANT: Ruvkun, Gary
; APPLICANT: Kimura, Koutarou
; APPLICANT: Patterson, Garth
; APPLICANT: Ogg, Scott
; APPLICANT: Paradis, Suzanne
; APPLICANT: Tissenbaum, Heidi
; APPLICANT: Morris, Jason
; APPLICANT: Kowsek, Allison
; TITLE OF INVENTION: THERAPEUTIC AND DIAGNOSTIC TOOLS FOR
; FILE OF INVENTION: IMPAIRED GLUCOSE TOLERANCE CONDITIONS
; CURRENT APPLICATION NUMBER: US/09/844,353A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 08/857,076
; PRIOR FILING DATE: 1997-05-15
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-844-353A-45

Query Match 6.0%; Score 73; DB 10; Length 510;
Best Local Similarity 20.9%; Pred. No. 18;
Matches 48; Conservative 37; Mismatches 11; Indels 34; Gaps 9;

Qy 4 SEEINEKDLRKSELQGTALGNLKOIYYNSKAITSSSEKSAQDQFLNTLLFKGFTGHPW 63
Db 125 SESPDVTVSGKKTTRRNAWGNMSYAEILTTAIMASPEKR---LTLAQVYEMWQVNPY 180

Qy 64 YNDLLVLGSTAANSEYEGSSVDLYGAYGQCAGGTPNKACMYGGVTLHDNNRLTEEK 123
Db 181 FRD-KGDSNSSAGMKNSIRHNLSLHSRPMRIQNEGAGKSSWWVINPDAPGMNPRTRER 239

Qy 124 KVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHLYHGKFGLYNSDSFGGKVQRGL 183
Db 240 SNTIE-----TTTQAQLEKSRGAKK-RIKERALMGS--LHSTL---NGNSIAGSIQ--- 285

Qy 184 IVFHSSEGSTVSYDLFD---AQGQY---PDTLLRIRYDNTTISSTLSIS 227
Db 286 -----TISHDLYDDDSMQGAFDNVPSSFRPRTQSNLSPGSSSRVS 326

RESULT 41
US-09-801-368-418
; Sequence 418, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Royer, John
```

```
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 418
; LENGTH: 566
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-418

Query Match 6.0%; Score 73; DB 10; Length 566;
Best Local Similarity 20.5%; Pred. No. 21;
Matches 38; Conservative 31; Mismatches 50; Indels 66; Gaps 8;

Qy 7 INEKDLRKSELQGTALGNLKOIYYNSK-----AITSSEKSAQDQFLNTLLFL 54
Db 295 VNGVDMRKA-----SSHEYNNVEVPFETKLSRIGISSLEKSRNQLLNDILF 343

Qy 55 KGFTTGHWPYNDLLVDLG-----STAATSEYEGSSVDLYGAYGQCAGGTPNKACMYG 109
Db 344 -----NNALTDLNLKSLATSIERLYKNSNITM-----KTKT----- 375

Qy 110 GVTLHDNNRLTEEKVVPINLWIDGKQTTVPIDKVKTSKKEVTVQELDLQARHLYHGKFG 169
Db 376 -LNQDESTVNEPKT-----RRPLLIDREKFLNKEFLDEIAIEIYFTLSEPKD 425

Qy 170 YNSDS 174
Db 426 LNSDT 430

RESULT 42
US-10-014-436-4
; Sequence 4, Application US/10014436
; Publication No. US20020182699A1
; GENERAL INFORMATION:
; APPLICANT: HATADA, Yuji
; APPLICANT: IGARASHI, Kazuaki
; APPLICANT: OZAKI, Katsuya
; APPLICANT: ARA, Katsutoshi
; APPLICANT: KAWAI, Shuji
; APPLICANT: ITO, Susumu
; TITLE OF INVENTION: Gene for Enzyme Having Both Alkaline Pullulanase and Alkaline
; FILE OF INVENTION: Amylase Activities
; FILE REFERENCE: 2173-0122P
; CURRENT APPLICATION NUMBER: US/10/014,436
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: JP 111547/1995
; PRIOR FILING DATE: 1995-05-10
; PRIOR APPLICATION NUMBER: PCT/JP96/01243
; PRIOR FILING DATE: 1996-05-10
; PRIOR APPLICATION NUMBER: US 08/952,084
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: US 09/514,302
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 893
; TYPE: PRT
; ORGANISM: Bacillus sp.
US-10-014-436-4
```

[illegible]

```

RESULT 43
US-08-837-459-19
; Sequence 19, Application US/08837459
; Patent No. US2002006407A1
; GENERAL INFORMATION:
; APPLICANT: McKee, Marian L.
; APPLICANT: O'Brien, Allison D.
; APPLICANT: Wachtel, Marian R.
; TITLE OF INVENTION: Histidine-Tagged Intimin And Methods Of
; TITLE OF INVENTION: Using Intimin To Stimulate An Immune Response And As An
; TITLE OF INVENTION: Antigen Carrier With Targeting Capability.
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/837,459
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fordis, Jean B.
; REGISTRATION NUMBER: 32,984
; REFERENCE/DOCKET NUMBER: 04995.0023-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4400
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 934 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-837-459-19

```

Query Match 6.0%; Score 73; DB 8; Length 934;
Best Local Similarity 21.4%; Pred. No. 41;
Matches 47; Conservative 32; Mismatches 77; Indels 64; Gaps 11;

```

Qy 21 TALGNLKOIYYNNSKAITSSKSDAQFLNTNLLPKGFFTHGPWYNLDLVDLGSAAITSEY 80
   || : || : || : || : || : || : || : || : || : || : || : || : || :
Db 669 TAVANGKDAIKYTVKVMKNGQPVNNQSVTFSTNF-GMFNGK-----SQTOATTGN 717

Qy 81 EG-SSVDLYGAYCYQCAGGTPNKATCMYG-GVTLHDNNRLTEEEKVPI-----N 128
   || : || : || : || : || : || : || : || : || : || : || : || : || :
Db 718 DGRATITLTSSAGKATVTSVSDGAEVKATEVTFDELKI--DNKVDIIIGNNVRGELPN 775

Qy 129 LWIDGKQTTVPIDRKVKS-----KKEVTVGELDLOARHYLHGKFGLYNDSFGKGV 179
   || : || : || : || : || : || : || : || : || : || : || : || : || :
Db 776 IWLYGQF-----KLKASGGDGTYSWTSNTSIAVDASGKVTLNGK----- 817

Qy 180 QRLGILVHFSSEG--STVSYDL-----FDAQGQYPTDL 209
   || : || : || : || : || : || : || : || : || : || : || : || : || :
Db 818 --GSVVIKATSGDKQTSYTIKAPSYMIKVDPKOAYYADAM 855

```

```

RESULT 44
US-09-801-368-304
; Sequence 304, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 304
; LENGTH: 1155
; TYPE: prt
; ORGANISM: Saccharomyces cerevisiae
; US-09-801-368-304

```

Query Match	6.0%; Score 73; DB 10;	Length 1155;
Best Local Similarity	20.6%; pred. No. 55;	
Matches 53; Conservative 34; Mismatches 84; Indels 86; Gaps 11;		
Qy	26 LKQIYYVNSKAITSEKSADOLFNTLLFKGFFTHGPWYNLLVDLGSTAATSEYEGSSV	85
	: : : : : :	
Db	147 LPHVRKHSSPLLSEFSKN-----GSHMGDP--NQL-----STPPTPKSAGHTM	188
Qy	86 DLYGAYGYQCAGTPN-----KT-----ACMYGGVTLHNNRLTEKKVPKI	127
	: : : : : :	
Db	189 ELHSFNGKHSSSTSLSFALESUKTNRRSNSNHSSQYRRTNQHRHRSKSPV	248
Qy	128 NL----WDTGQTIVP-----IDKVTKSKEVTVQELDQAARHLHGKFGLYNDSDEGG	177
	: : : : : :	
Db	249 SLTEISMIGTPLYVPALLSLIAIKFKOTIKLST-----HKMGGLYRDSFTG	296
Qy	178 KV-----QRGLVFISSEGSTVSVDLFDAQOYPDTLLRIY-----	213
	: : : : : :	
Db	297 KOAIDTCLLTIGSDLRNUGMLIGKLSLEAKQLFDHVDLYDHG--VROSVLLEYELSESIFMA	355
Qy	214 ---RDNTTISSTISLSIS	227

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:55:11 ; Search time 8.75359 Seconds
(without alignments)
2558.872 Million cell updates/sec

Title: US-09-900-766-2
Perfect score: 1218
Sequence: 1 SEKSEINEKDLRKSELOQ.....RDNTTISSTLSLSLYLTT 233

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1107	90.9	257	2	enterotoxin E prec
2	948	77.8	257	2	enterotoxin A prec
3	918	75.4	260	2	enterotoxin P [imp
4	611	50.2	258	2	enterotoxin D prec
5	420.5	34.5	258	2	enterotoxin Sen [i
6	372.5	30.6	260	2	enterotoxin SeO [i
7	311	25.5	240	2	extracellular ente
8	296.5	24.3	242	2	extracellular ente
9	284.5	23.4	266	2	enterotoxin C3 - S
10	282	23.2	239	2	enterotoxin SEM [i
11	277.5	22.8	266	2	enterotoxin C-2 pr
12	269.5	22.1	251	1	exotoxin type A pr
13	268.5	22.0	266	1	enterotoxin B prec
14	263.5	21.6	266	1	enterotoxin C-1 pr
15	253.5	20.8	236	2	exotoxin type A pr
16	251.5	20.6	236	2	exotoxin A precurs
17	248.5	20.4	236	2	exotoxin type A pr
18	241	19.8	258	2	extracellular ente
19	220	18.1	250	1	streptococcal pyro
20	190	15.6	136	2	enterotoxin YENT2
21	178	14.6	157	2	hypothetical prote
22	155.5	12.8	235	2	exotoxin C precurs
23	115	9.4	62	2	hypothetical prote
24	100	8.2	234	1	toxic shock syndro
25	98	8.0	133	2	enterotoxin Yent1
26	98	8.0	227	2	exotoxin 14 [impor
27	98	8.0	234	2	toxic shock syndro
28	96	7.9	552	2	chitinase chi-A or
29	94	7.7	231	2	exotoxin 11 [impor

RESULT 1

A28179

enterotoxin E precursor - Staphylococcus aureus

C:Species: Staphylococcus aureus

C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 15-Oct-1999

C:Accession: A28179

R:Couch, J.L.; Soltis, M.T.; Betley, M.J.

J. Bacteriol. 170, 2954-2960, 1988

A:Title: Cloning and nucleotide sequence of the type E staphylococcal enterotoxin gen

A:Reference number: A28179; MUID:88257005; PMID:3384800

A:Accession: A28179

A:Molecule type: DNA

A:Residues: 1-257 <COU>

A:Cross-references: GB:M21319; NID:gl53001; PIDN:AAA26617.1; PID:gl53002

C:Superfamily: enterotoxin B

Query Match 90.9%; Score 1107; DB 2; Length 257;
Best Local Similarity 89.7%; Pred. No. 9.8e-88;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELOQTALGNLKQIYYYNKRAITSEKSADQFLNTLFLKGFFTG 60
|||||
Db 25 SEKSEINEKDLRKSELOQNALSNLRQIYYNEKAITENKESDDQFLENTLFLKGFFTG 84
QY 61 HPWYNDLLVDLGSTATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
|||||
Db 85 HPWYNDLLVDLGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWIDGKQTPVPIDKVKTSKEVTVOELDQARHYLHGKFGLYNSDSFGGKVQ 180
|||||
Db 145 EEKVPINLWIDGKQTPVPIDKVKTSKEVTVOELDQARHYLHGKFGLYNSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTSVSYDLFDAQOQYPTDLLRIYRDNTTISSTLSLSLYLTT 233
|||||
Db 205 RGLIVFHSSEGSTSVSYDLFDAQOQYPTDLLRIYRDNTTINSENLDLYLTT 257

RESULT 2

A28664

enterotoxin A precursor - Staphylococcus aureus

C:Species: Staphylococcus aureus

C>Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 15-Oct-1999

C:Accession: A28664; A29566

R:Betley, M.J.; Mekalanos, J.J.

J. Bacteriol. 170, 34-41, 1988

A:Title: Nucleotide sequence of the type A staphylococcal enterotoxin gene.

A:Reference number: A28664; MUID:88086892; PMID:3335483

A:Accession: A28664

A:Molecule type: DNA

A:Residues: 1-257 <BET>

A:Cross-references: GB:M18970; NID:gl53120; PIDN:AAA26681.1; PID:gl53121

A:Experimental source: strain FRI337

Pl15 protein - Myc
outer membrane sec
DNA-directed DNA p
thymidine phosphor
pheromone CAD1 bin
hypothetical prote
hypothetical prote
probable myrosinas
myrosinase-binding
hypothetical prote
glucosyltransferas
chromodomain helic
putaive autotransp
glucosyltransferas
glycosidase homolo
outer membrane por

Db 92 SSLKVEFNSSDLANOFKGNKNDIYGLYFGNCKVGLTEKTSCLYGVTTIHDGNOLDEKV 151
QY 125 VPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFLYNSDFSFGKQVORGLI 184
Db 152 IGWVFKDGVQOEGFV--IKTKAKVTVOELDTKVRKLENIYKIYNKDT--GNIOKGCI 207
QY 185 VFHS--SEGTSVSYDLFDAQGQYPTDRLRIYRDNTTISSTLSLSILYLY 231
Db 208 FFHSHNHQDSFYDLYNKGSGVGAFFQFYSDNRTVSSSNHYHIDVFLY 256

RESULT 6
E89969
enterotoxin seO [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: E89969
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: E89969
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-260 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701623; PIDN:BAB42916.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: seO
C:Superfamily: enterotoxin B

Query Match 30.6%; Score 372.5; DB 2; Length 260;
Best Local Similarity 38.8%; Pred. No. 1.5e-24;
Matches 93; Conservative 32; Mismatches 92; Indels 23; Gaps 7;
QY 8 NEKD-----LRKKSLEQGTALGNLKOIYVNS-KAITSSEKSAQDQFLNTLLFKGFFTG 61
Db 29 NEEDPKIESLCKSSVDPTALHNDYNNRFTTVKSVSTTEKFLDPDLLFKSI----- 84
QY 62 PW-----YNDLLDGLSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMGVTLH 114
Db 85 NMLDGTSAEFKDLKVEFSSAISKEFLGKTVDIYGVYKHAHGEHQVDVDTACTYGGVTPH 144
QY 115 DNNRLTEKKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFLYNSDS 174
Db 145 ENKUSEPKNIGVAVYKONVNTFI--VTTDKKVTQAQELDIKVRTKLNNAYKLY--DR 200
QY 175 FGGKVQRGVLIVFHSSEGSTVS--YDLFDAQGQYPTDRLRIYRDNTTISSTLSLSILYLT 232
Db 201 MTSQVQKGVYKFSHSEKHSKESFYDLFIYKGNLPDQYLQIYNDKNTIYDSSDHYDVLFT 260

RESULT 7
G89991
extracellular enterotoxin L [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89991
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G89991
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-240 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701803; PIDN:BAB43096.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:

A:Gene: sel
C:Superfamily: enterotoxin B

Query Match 25.5%; Score 311; DB 2; Length 240;
Best Local Similarity 33.3%; Pred. No. 2.6e-19;
Matches 74; Conservative 36; Mismatches 78; Indels 34; Gaps 6;
QY 24 GNLKQIY----YNSKAITTSSEKSAQDQFLNTLLFKGFFTGHPWYND--LLVDLGSTAATS 78
Db 31 GNLRFYTKYEYVNLKNVKNKNSPESHRL-----YSYKNDTLAEFDNEYITS 79
QY 79 EYEGSSVDLYGAYGYOCAGGTPNKTACMGVTLHDNNRLTEKKVPINLWIDGKQTTV 138
Db 80 DLKGNVDVFGISYKY----GNSRT--IYGGVTRAEKNKLDSPRIPIINLIINGKHQTV 133
QY 139 PIDKVKTSKEVTVQELDLQARHYLHGKFLY-----NSDSFGGKVGQRLIVF 186
Db 134 TPKSVSTDKMVTQAQEIYDKLRYLQDEFNIIYGHNDTGKKEYGTSKPFYSGFDKGSVVF 193
QY 187 HSEGSTSVSYDLFDAQGQYPTDRLRIYRDNTTISSTLSLSISL 228
Db 194 HMDGNSFSYDLFYTGYGLPESFLKIYKDNKTVDSTQPHLDV 235

RESULT 8
C89969
extracellular enterotoxin type I precursor [Imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C89969
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C89969
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-242 <KUR>
A:Cross-references: GB:BA000018; PID:gl3701621; PIDN:BAB42914.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: sei
C:Superfamily: enterotoxin B

Query Match 24.3%; Score 296.5; DB 2; Length 242;
Best Local Similarity 32.4%; Pred. No. 4.7e-18;
Matches 77; Conservative 39; Mismatches 85; Indels 37; Gaps 10;
QY 8 NEKDLRKSELOG-TALGNLKOIY----YNSKAITTSSEKSAQDQFLNTLLFKGFFTGHP 62
Db 17 NIKDL---TYAAGDIGVGNLRFNFTKHDYIDLKGVTDKNLP----IANQLEFS---TG-- 64
QY 63 WYNDLLVDLGSTAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMGVTVLHDNNRLTEE 122
Db 65 -TNDLISNSNWEDESKFKGKKLDIFGIDYNGPC-----KSKYMYGGATL-SGOYLNSA 116
QY 123 KKVPIINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFLYNSDSFG--GK--- 178
Db 117 RKIPINLWVNGKHKTISTDKIATNKKLTQAQEIYDKLRYLQEEYNIYGHNTGKKEYG 176
QY 179 -----VQRLIVFHSSEGSTVSYDLFDAQGQYPTDRLRIYRDNTTISSTLSLSISL 228
Db 177 YKSKEYFGNNGKVLFLHNLNNEKFSYDLFYTGDGGLPVSFLKIYEDNKITIESEKPHLDV 234

RESULT 9
S11885
enterotoxin C3 - Staphylococcus aureus
C:Species: Staphylococcus aureus
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Jul-1999
C:Accession: S11885

R:Hovde, C.J.; Hackett, S.P.; Bohach, G.A.

Mol. Gen. Genet. 220, 329-333, 1990
A:Title: Nucleotide sequence of the staphylococcal enterotoxin C3 gene: sequence comparison
A:Reference number: S11885; MUID:90220508; PMID:2325627

A:Accession: S11885
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-266 <HGV>
A:Cross-references: GB:X51661; NID:g46570; PIDN:CAA35972.1; PID:g46571
C:Superfamily: enterotoxin B

Query Match 23.4%; Score 284.5; DB 2; Length 266;
Best Local Similarity 32.5%; Pred. No. 5.8e-17;
Matches 76; Conservative 45; Mismatches 94; Indels 19; Gaps 9;

Qy 11 DLRLKSELQGTALGNLKOIYYNSKAITSS-ESADQFLNTLLFKGFTGHPWYNDLLV 69
Db 37 DLHKSSETGT-MGNMK--YLYDDHYVSATKVKSVDFLAHDLIYNISDKLNKYDKVKT 93
Qy 70 DLGTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGGVTLHDNNRLTEE 122
Db 94 ELLNEDLAKKYDEVVDYGSNYVNCYFSSKDNVGVTKGTCMYGGITKHEGHPDNG 153
Qy 123 --KKVPINLWIDGKQTTVPIDKVKTSKEVTQVQLDQARHYLHGKFGVSGGKQV 180
Db 154 NLQNVLRVY-ENKRNITISFE-VQTDKKSVAQELDIKARNFLINKKNLYEFNS--SPYE 209
Qy 181 RGLIVFHSSEGSTVSYDLFDAQGYPD--TLRIYRDNTTISSTSLSLYLTY 232
Db 210 TGYIKFIENNGTWFYDMMMPAGDKFDOSKYLMMYNDKNTVDSKSVKIEVHLT 263

RESULT 10
D89969
enterotoxin SEM [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.: Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758; MUID:21311952; PMID:11418146

A:Accession: D89969
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-239 <KUR>
A:Cross-references: GB:BA000018; PID:g13701622; PIDN:BA042915.1; GSPDB:GN00149
A:Experimental source: strain N315

C:Genetics:
A:Gene: sem
C:Superfamily: enterotoxin B

Query Match 23.2%; Score 282; DB 2; Length 239;
Best Local Similarity 29.7%; Pred. No. 8.2e-17;
Matches 68; Conservative 40; Mismatches 77; Indels 44; Gaps 6;

Qy 23 LGNLKOIYYNSKAIT--SSEKADQFLNTLLFKGFTGHPWYNDLLVDLGSTAAATSE- 79
Db 24 VGVNLNRYGSPYEDHQSPENNNHLSQLVF-----SMDNSTVTAEF 68
Qy 80 -----YEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWI 131
Db 69 KNVDVKKFKHADVYGLSVGYCL-----KNKYIYGGVTL-AGDYLEKSRPPIINLWV 122
Qy 132 DGKQTTVPIDKVKTSKEVTQVQLDQARHYLHGKFGV-----NSDSFGGKV 179
Db 123 NGEHQTISTDKVSTNKKLVTAQETDKLRRYLQBEYNIYFNDTKGRNYGNKSKFSSGF 182
Qy 180 QRLIVFHSSEGSTVSYDLFDAQGYPD--TLRIYRDNTTISSTSLSL 228
Db 183 NAGKILFLNDGSSFSYDLFTGTGTGQAEFLKIYNDKNTVETKFLDV 231

RESULT 11

A60114
enterotoxin C-2 precursor - Staphylococcus aureus
N:Alternate names: enterotoxin C-3 precursor
C:Species: Staphylococcus aureus
C:Date: 10-Nov-1992 #sequence_revision 10-Nov-1992 #text_change 16-Jul-1999
C:Accession: A60114; B60114; A33866
R:Bohach, G.A.; Schlievert, P.M.
Infect. Immun. 57, 2249-2252, 1989

A:Title: Conservation of the biologically active portions of staphylococcal enterotoxin
A:Reference number: A60114; MUID:89277549; PMID:2543637

A:Accession: A60114
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-266 <BOH>
A:Accession: B60114
A:Molecule type: protein
A:Residues: 28-66 <BOH2>
R:Couch, J.L.; Betley, M.J.
J. Bacteriol. 171, 4507-4510, 1989

A:Title: Nucleotide sequence of the type C-3 staphylococcal enterotoxin gene suggests
A:Reference number: A33866; MUID:89327174; PMID:2473979

A:Accession: A33866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-266 <COU>

A:Cross-references: GB:M28364; NID:g153003; PIDN:AAA26624.1; PID:g153004
C:Genetics:
A:Gene: entC2

C:Superfamily: enterotoxin B
F:1-27/Domain: signal sequence #status predicted <SIG>
F:28-266/Product: enterotoxin C-2 #status predicted <MAT>

Query Match 22.8%; Score 277.5; DB 2; Length 266;
Best Local Similarity 31.1%; Pred. No. 2.3e-16;
Matches 76; Conservative 47; Mismatches 102; Indels 19; Gaps 9;

Qy 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNSKAITSS-ESADQFLNTLLFKGFT 59
Db 27 AESQPDPTDELHKSSSEFTGT-MGNMK--YLYDDHYVSATKVKSVDFLAHDLIYNISDK 83
Qy 60 GHPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGVT 112
Db 84 KLVNVDKVKTELLNEDLAKKYDEVVDYGSNYVNCYFSSKDNVGVTKGTCMYGGIT 143
Qy 113 LHDNNRLTEE--KKVPINLWIDGKQTTVPIDKVKTSKEVTQVQLDQARHYLHGKFGV 170
Db 144 KHEGHPDNGNLQNVLRVY-ENKRNITISFE-VQTDKKSVAQELDIKARNFLINKKNLY 201
Qy 171 NSDSFGGKVQRLIVFHSSEGSTVSYDLFDAQGYPD--TLRIYRDNTTISSTSLSL 228
Db 202 EFNS--SPYETGYIKFIENNGTWFYDMMMPAGDKFDOSKYLMMYNDKNTVDSKSVKIEV 259
Qy 229 YLYT 232
Db 260 HLT 263

RESULT 12

S29659
exotoxin type A precursor (allele 1) - Streptococcus pyogenes phase T12
N:Alternate names: erythrogenic toxin; scarlet fever toxin
C:Species: Streptococcus pyogenes phase T12
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S29659; S18782; S18784; S18785; S18791; S18796; S18800
R:Weeks, C.R.; Ferretti, J.J.
Infect. Immun. 52, 144-150, 1986

A:Title: Nucleotide sequence of the type A streptococcal exotoxin (erythrogenic toxin
A:Reference number: S29659; MUID:8616804; PMID:3514452

A:Accession: S29659
A:Molecule type: DNA

C:Species: Streptococcus pyogenes
A:Variety: strain MGAS262 isolate California
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S18789
R:Nelson, K.; Schlievert, P.M.; Selander, R.K.; Musser, J.M.
J. Exp. Med. 174, 1271-1274, 1991
A:Title: Characterization and clonal distribution of four alleles of the speA gene encoded by the enterotoxin A precursor (allele 2) - Streptococcus pyogenes phase (strain MGAS250 isolate N18786)
A:Reference number: S18788; PMID:92044323; PMID:1940804
A:Accession: S18789
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61573; NID:g47303; PIDN:CAA43771.1; PID:g47304
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
C:Genetics:
A:Gene: speA
C:Superfamily: enterotoxin B
C:Keywords: exotoxin
F:1-22/Domain: signal sequence (fragment) #status predicted <SIG>
F:23-236/Product: exotoxin A (fragment) #status predicted <MAT>

Query Match 20.6%; Score 251.5; DB 2; Length 236;
Best Local Similarity 31.7%; Pred No. 3.4e-14;
Matches 72; Conservative 42; Mismatches 102; Indels 11; Gaps 7;

QY 4 SEINEKDLRKSELOGTAL-GNLKOIY--YNSKAITSSSEKSAQDLTNTLLFKGFFTG 60
Db 17 SQEVFAQQDPNPQLSRSLVKNLQNIYFLYEGDPVHENVKSVQDQLSHDLIYN--VS 73

QY 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQC-AGTPNKTKACMYGGVTLHDNNRL 119
Db 74 GLNYDKLKTTELKQEMATLFLKDKNDVIYVEYHYLCYLCENAEPSACIYGGVTNHEGNHL 133

QY 120 TEEKKVPINLWDGKQTTVPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKV 179
Db 134 EIPKNTLVKVSIDGQ-SLSFD-IETSKMVTAEQLDYKVRKHLTDNKKQLYNGP--SKY 189

QY 180 QRLGIVFHSSEGSTVSYDLFDAGQGYPDTLRLIYRDNTTISSTLSI 226
Db 190 ETGYIKFISKDKETWFDFPEPEFQVKNVLYMKDNETLDSSTSOI 236

RESULT 17
S18786
exotoxin type A precursor (allele 2) - Streptococcus pyogenes phase (strain MGAS250 isolate N18786)
N:Alternate names: scarlet fever toxin
C:Species: Streptococcus pyogenes phase
A:Variety: strain MGAS250 isolate California; strain MGAS251 isolate California; strain MGAS252 isolate California
C:Date: 29-Jan-1993 #sequence_revision 29-Jan-1993 #text_change 16-Jul-1999
C:Accession: S18786; S18787; S18788; S18790; S18792; S18795; S18799
R:Nelson, K.; Schlievert, P.M.; Selander, R.K.; Musser, J.M.
J. Exp. Med. 174, 1271-1274, 1991
A:Title: Characterization and clonal distribution of four alleles of the speA gene encoded by the enterotoxin A precursor (allele 2) - Streptococcus pyogenes phase (strain MGAS250 isolate N18786)
A:Reference number: S18788; PMID:92044323; PMID:1940804
A:Accession: S18786
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61561; NID:g47297; PIDN:CAA43759.1; PID:g47298
A:Experimental source: strain MGAS250 isolate California unassigned phase
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18787
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61562; NID:g47299; PIDN:CAA43760.1; PID:g47300
A:Experimental source: strain MGAS251 isolate California unassigned phase
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18788
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>

A:Cross-references: EMBL:X61563; NID:g47301; PIDN:CAA43761.1; PID:g47302
A:Experimental source: strain MGAS256 isolate California unassigned phase
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18790
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61564; NID:g47305; PIDN:CAA43762.1; PID:g47306
A:Experimental source: strain MGAS285 isolate Colorado unassigned phase
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18792
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61565; NID:g47311; PIDN:CAA43763.1; PID:g47312
A:Experimental source: strain MGAS480 isolate Yugoslavia unassigned phase
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18795
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61566; NID:g47317; PIDN:CAA43764.1; PID:g47318
A:Experimental source: strain MGAS492 isolate United Kingdom unassigned phase
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
A:Accession: S18799
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-236 <NE>
A:Cross-references: EMBL:X61567; NID:g47325; PIDN:CAA43765.1; PID:g47326
A:Experimental source: strain MGAS496 isolate Germany unassigned phase
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, September 1991
C:Genetics:
A:Gene: speA2
C:Superfamily: enterotoxin B
C:Keywords: exotoxin
F:1-22/Domain: signal sequence (fragment) #status predicted <SIG>
F:23-236/Product: exotoxin type A (fragment) #status predicted <MAT>

Query Match 20.4%; Score 248.5; DB 2; Length 236;
Best Local Similarity 31.6%; Pred No. 6.1e-14;
Matches 73; Conservative 44; Mismatches 95; Indels 19; Gaps 10;

QY 4 SEINEKDLRKSELOGTAL-GNLKOIY-YNSKAITT-SSEKSAQDLTNTLLFKGFFTG 60
Db 17 SQEVFAQQDPNPQLSRSLVKNLQNIYFLYEGDPVHENVKSVQDQLSHDLIYN--VS 73

QY 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQC-AGTPNKTKACMYGGVTLHDNNRL 119
Db 74 GPNYDKLKTTELKQEMATLFLKDKNDVIYVEYHYLCYLCENAEPSACIYGGVTNHEGNHL 133

QY 120 TEEKKVPINLWDGKQTTVPIDKVTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGGKV 179
Db 134 EIPKNTLVKVSIDGQ-SLSFD-IETSKMVTAEQLDYKVRKHLTDNKKQLYNGP--SKY 189

QY 180 QRLGIVFHSSEGSTVSYDLFD---AQGYPDTLRLIYRDNTTISSTLSI 226
Db 190 ETGYIKFIPKNESEWFDFPEPEFTQSKY----LMIYKDNETLDSNTSOI 236

RESULT 18
G89968
extracellular enterotoxin type G precursor [imported] - Staphylococcus aureus (strain C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89968
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oma, A.; Mizutani-Uti, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; PMID:21311952; PMID:11418146
A:Accession: G89968
A:Status: preliminary

A: Molecule type: DNA
A: Residues: 1-258 <KUR>
A: Cross-references: GB:BA000018; PID: gl3701617; PID: BAB42910.1; GSPDB: GN00149
A: Experimental source: strain N315
C: Geneticals:
C: Gene: seg
C: Superfamily: enterotoxin B

```

Query Match      19.8%; Score 241; DB 2; Length 258;
Best Local Similarity 28.8%; Pred. No. 3e-13;
Matches 72; Conservative 48; Mismatches 88; Indels 42; Gaps 12;

Qy      3 KSEINEKDLRKKBELOGTALGNLKKQIYY---YNSKAITTSSEKSADQFLTNTLLFKGFFT 59
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      30 KLDELKVKSDYKNN--KGT-MGNVMNLITSPPVEGRGVINSR-----QFLSHDLIFP---I 79

Qy      60 GHPWYNLDLVLGSTATSEYEGGSDVLYGAYGYQCA-----GTPNKKTACM 107
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      80 EYKSYNEVKTELENTELANNYKDKVDIFGVPYFTCIPIKSEPDINQFVG-----CCM 134

Qy      108 YGGVTLH--DNNRLTEEEKKVPINLWIDGKQTVTPIDKVKTSKEVTVQELDLQARHYLHG 165
      | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      135 YGGLTFNSENER--DKLITVQVTIDNRQSIG--FTITTNNMVTIQELDYKARHWLTK 189

Qy      166 KFLGYNSDSFGCKVQRGILVPHSSEGSTVSDYDLFDAQQQYPD-----LRIYRDNTTIST 222
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      190 EKKIYEFD--GSAFESGVYIKFTEKNNTSFWFDLPKELVPFVFKFLNIYDKNKVVSD 247

Qy      223 SLISSLYLT 232

Db      248 SIKMEVFLNT 257

```

RESULT 19
A26152
streptococcal pyrogenic exotoxin type A precursor - Streptococcus sp.
N:Alternate names: scarlet fever toxin; SPE type A (speA)
C:Species: Streptococcus sp.
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: A26152
R:Johnson, L.P.; L'Italien, J.J.; Schlievert, P.M.
Mol. Gen. Genet. 203, 354-356, 1986
A:Title: Streptococcal pyrogenic exotoxin type A (scarlet fever toxin) is related to Staphylococcus aureus exfoliative toxin A
A:Reference number: A26152; MUID:86284313; PMID:3526093
A:Accession: A26152
A:Molecule type: DNA
A:Residues: 1-250 <JOH>
C:Superfamily: enterotoxin B
C:Keywords: exotoxin

Query Match	18.1%	Score 220;	DB 1;	Length 250;
Best Local Similarity	28.9%	Pred. No. 1.9e-11;		
Matches 66; Conservative 42; Mismatches 98; Indels 22; Gaps 9;				
Qy	14	KKSELOGTALGNLKOIY-YYNKSAITSE--KSADQFLTNLLPKGFTTGPWYNLDLVD	70	
Db	35	KPSQLQRSLNVKTKIYFFRVRVLVTHENVKSVQDLLSHDLIYN--VSGPNDKCLKTE	91	
Qy	71	LGSTAAATSEYEGSSVDLYGAYGYQC-AGGTPNTACMGYGVTLHDNNRLRTEKKVPINL	129	
Db	92	LKNQEMATLFKDKKNDVYGVGYHCLYLCENARSACLYGGVTNHEGHLEIPKKIVVKV	151	
Qy	130	WIDGKQT-TVPIDKRVTSKKEVTQOEIQLQARHVLHGKFGLYNDSFGGKVGORGLIVFHS	188	
Db	152	SIDGIQSLSPDIEQIKNG---NCSRISYTVRKYLTDNKQLYTNGP--SKYETGYKEIP	205	
Qy	189	SEGTSVSVLDLED---AQGVDPDILLRIYRDNNTTISSTLSISLYLYT	232	
Db	206	KNKESFFWDFPEPEFTQSKY---LMIYKDNETLDSNTSQIEVLYLT	249	

RESULT 20
A89969

enterotoxin YEW2 [Imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
R:Accession: A89969
C:Curator: M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; O
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaiko, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357. 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: A89969
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-136 <KUR>
A:Cross-references: GB:BA000018; PID:g13701619; PIDN:BAB42912.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: vent2

Query Match	15.6%	Score 190;	DB 2;	Length 136;
Best Local Similarity	34.1%;	Pred. No. 3.2e-09;		
Matches	46;	Conservative 25;	Mismatches 54;	Indels 10; Gaps 5;
Qy	103	KTACMYGVTTHDNNRLTEKKV--PINLWI--DGKQTTPIDIKVKTSKREVTVOELDL	157	
		: : : : : : : : : : : : : : :		
Dd	2	KKTCMYGGVTHDGNQIDKNNSTDNHNLIVKENRNSLSFD--IPTNKKKNITAEIDY	60	
		: : : : : : : : : : : : : : :		
Qy	158	QARHYLHGKFGLYNDSFGGVQRGLIVFHSSEGSTVSVDLFDAGQG--YPDTLRLIYYRD	215	
		: :		
Dd	61	KVRNYLLKHKNLYEENS--SPYEICYIKFIHGSGHSFWYDLMPESGKKAFYPTKYLIIND	118	
		: : : : : : : : : : : : : : :		
Qy	216	NNTTSSSTSLSLYL	230	
		: :		
Dd	119	NKTVESKSINVEVHL	133	

RESULT 21

A89942

hypothetical protein SA1430 [imported] - Staphylococcus aureus (strain N315)

C:Species: Staphylococcus aureus

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001

C:Accession: A89942

R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Seki, C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K. Lancet 357, 1225-1240, 2001

A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A:Reference number: A89758, PMID:21311952, PMID:11418146

A:Accession: A89942

A>Status: preliminary

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 <KUR>

A:Cross-references: GB:BA000018; PID:g13701400; PID:BA042694.1; GSPDB:GN00149

A:Experimental source: strain N315

C:Genetics:

A:Gene: SA1430

```

Query Match          14.6%; Score 178; DB 2; Length 157;
Best Local Similarity 31.4%; Pred. No. 4.2e-08;
Matches 37; Conservative 21; Mismatches 58; Indels 2; Gaps 1;

QY 1 SEKSEINEKDLRKSELTQGTALGNLKIYYIYNKAITSSSEKSADQFNTLLPKGFFTG 60
   : : | | | | : : : : : : : : : : : : : : : :
Db 42 TNSAIEYSDLHHKSKPDRSLNAKMSFINPTQ--LENNKNTDRLLKHDLFLPHDFMEVN 99
   : : : : : : : : : : : : : : : : : : : : : :

QY 61 HPWINDLLVLGSLTAATSEYEGSSVDLYGAYGYOCAGGTENKTKACMYGGVTLHDNNR 118
   | : : : : : : : : : : : : : : : : : : : : :
Db 100 VASKDKFKEFENEALSKKFINKIDIVAGSYSECHGGATNKQCSGGVTLSDNNK 157
   : : : : : : : : : : : : : : : : : : : : :

RESULT 22
A30509
exotoxin C precursor - Streptococcus pyogenes

```



```

RESULT 26
G89807
exotoxin 14 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: G89807
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: G89807
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-227 <KUR>
A:Cross-references: GB:BA000018; PID:g13700320; PIDN:BAB41618.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: set14

Query Match 8.0%; Score 98; DB 2; Length 227;
Best Local Similarity 26.4%; Pred. No. 0.51;
Matches 53; Conservative 28; Mismatches 78; Indels 42; Gaps 9;

QY 8 NEKDRLKSELOGTALGNLKQIY-YNNSKAITSSKSDAOFNTNLLFKGFTGHPWYND 66
Db 33 NOKSVNKHDK-----EALRYITGKTMKMKNISALKHGKNNLRFK--FRGIKIQVL 81

QY 67 LLVDLGSTAATSEYEGSSV-----DLYGAYGYQCAGGT--PNKTACMGYGVTLHDNN 117
Db 82 LPGNDKSKFQORSYEGLDVFFVQERKHDIFY---TVGCVTONKNTS-----GVVSAPIL 134

QY 118 RUTEKVPINLWIDKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGG 177
Db 135 NISKEK-----GEDAFVKGYPYIKKEKITLKDLYDKRKHLIEKGYLKIYSKDG 185

QY 178 KVQGLIVFHSSEGSTVSYDL 198
Db 186 RVKISL-----KDGsfYNLDL 201

RESULT 27
B89992
toxic shock syndrome toxin-1 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89992
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89992
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <KUR>
A:Cross-references: GB:BA000018; PID:g13701806; PIDN:BAB43099.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: tst
C:Superfamily: toxic shock syndrome toxin

Query Match 8.0%; Score 98; DB 2; Length 234;
Best Local Similarity 24.6%; Pred. No. 0.54;
Matches 49; Conservative 33; Mismatches 57; Indels 60; Gaps 13;

QY 25 NLKQI---YYNSKAITSE-----KSADQFLTNLLFKGFTGHPWYNDLL-- 68
Db 45 NIKDLLDWSGSDFTNSELVNSLGMRIKNTDGS1-SLIIFPS-----PYSPAFTK 98
```

```

QY 69 ---VDLGSTAATSEYEGSSVDLYGAYGYQCAGGTGNKTACMGYGVTLHDNNRLTEKKV 125
Db 99 GEKVDL-NTRTKKKSQHTS---EGYIHFQISGVT-----NTEKLPTPIEL 140

QY 126 PINLWIDKQTTV---PIDKVKTSKKEVTVOELDLQARHYLHGKFGLY-NSDSFGG--KV 179
Db 141 PLKVKRVHGKDSPLKYWP-----NFDKKQLAISTLDLDFEIRHQLTQIHGLYRSSDKTGCVWKI 196

QY 180 QRGLIVFHSSEGSTVSYDL 198
Db 197 T-----MNDGSTYQSDL 208

RESULT 28
T41863
chitinase chi-A orf126 - Bombyx mori nuclear polyhedrosis virus (isolate T3)
C:Species: Bombyx mori nuclear polyhedrosis virus, BMSNPV
A:Variety: isolate T3
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
C:Accession: T41863
R:Gomi, S.; Wajima, K.; Maeda, S.
J. Gen. Virol. 80, 1323-1337, 1999
A:Title: Sequence analysis of the genome of Bombyx mori nucleopolyhedrovirus.
A:Reference number: Z22020; MUID:99281911; PMID:10355780
A:Accession: T41863
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-552 <KAM>
A:Cross-references: EMBL:L33180; NID:g3745835; PIDN:AAC63792.1; PID:g3745945
A:Experimental source: isolate T3
C:Superfamily: Serratia marcescens chitinase

Query Match 7.9%; Score 96; DB 2; Length 552;
Best Local Similarity 24.9%; Pred. No. 2.5;
Matches 46; Conservative 23; Mismatches 68; Indels 48; Gaps 11;

QY 32 YNSKAITS---SEKSADQFLTNLLFKGFTGHPWYNDLLVDLGSTAATSEYEGSSVDLY 88
Db 359 YDKIAVVNVAEAKSLDKIFLWYDFKG-----AWSN---TDLGYOTTVYVSWNSEELY 410

QY 89 GAYGYQC---AGGTPNKT---ACMYG-----GVTLDN-NRLTEKKVPIN-LWIDGKQT 136
Db 411 TTHYAVDALLEGQVDPNKIIVGVAMVGRGTGVTNTNGNYSFGTGNPVGSGTWEDG--- 467

QY 137 TVPIDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVGQRLIVFHSSEGSTVSY 196
Db 468 -----VVDYRQIQKDLNNY-----VYTFDS-----AAQASYVFDKSKGDLISF 505

QY 197 DLFLDA 201
Db 506 DSVDS 510

RESULT 29
D89807
exotoxin 11 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D89807
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of meticillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D89807
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <KUR>
A:Cross-references: GB:BA000018; PID:g13700317; PIDN:BAB41615.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
```


Query Match 7.5%; Score 91; DB 2; Length 1634;
Best Local Similarity 21.6%; Pred. No. 27;
Matches 50; Conservative 36; Mismatches 79; Indels 56; Gaps 10;

QY 12 LRKSELOCTALGNLQIYYNSKAITSEKSADQFLNTLLFKGFTTGHWPYNDLLVDL 71
Db 614 VRRKKGTRAITLGCCKADLYLKIEIL---KNKEKYLPAAIL-RGFFEDGQVYVRRRAV 669
QY 72 GSTAATSEYE-----GSSVDLYGAYYQCAGGTPNKTCACMYGGVTLHDNNRLTEKKVP 126
Db 670 VVNOGTNNYDKIFATSLDLRLGKYSF-----YTSYEERGGKLRKYV 713
QY 127 INLWIDGKQTPVIDKVKTS-----KKEVTVOELDLQARHYLHGKFGLYNSD----- 173
Db 714 IEIFSKG-----DLIKFSLISFISRRKNLLNEIIRKQTLKYGIDYGFYDLDDVCVS 766
QY 174 --SFGKV-----QRLGVHFSSEGS-TVSYDLFDQAQGYPTDL 209
Db 767 LESYKGEVVDLTLEGRPVYFANGILTHNSLYPSIIISYNI-----SPDTL 811

RESULT 33
C53312
thymidine phosphorylase (EC 2.4.2.4) - Mycoplasma pirum (strain BER)
C:Species: Mycoplasma pirum
C:Date: 23-Oct-1998 #sequence_revision 23-Oct-1998 #text_change 07-Dec-1999
C:Accession: C53312
R:Tham, T.N.; Ferris, S.; Kovacic, R.; Montagnier, L.; Blanchard, A.
J. Bacteriol. 175, 5281-5285, 1993
A:Title: Identification of Mycoplasma pirum genes involved in the salvage pathways for n
A:Reference number: A53312; MUID:93352438; PMID:8349569
A:Accession: C53312
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-419 <THA>
A:Cross-references: GB:L13289; NID:g401781; PIDN:AAA25432.1; PID:g401784
C:Genetics:
C:Superfamily: thymidine phosphorylase
C:Keywords: glycosyltransferase; pentosyltransferase

Query Match 7.2%; Score 88; DB 1; Length 419;
Best Local Similarity 27.0%; Pred. No. 8.3;
Matches 38; Conservative 22; Mismatches 57; Indels 24; Gaps 7;

QY 47 FLTNTLLFKG-FFTGHWPYNDLLVLGSGTAATSE-----YEGSSVDLYGAYYG 93
Db 57 FLTKAMIDSGKIYKFPPEYKTLIDKHSHTGGIGDKVSIALRPILVFDLGVAKLSGRGIG 116
QY 94 YOCAGGTPNKTCACMYGGVTLHDNNRLTEKKV--PINLWIDCK-QTTVPIDKVKTSKREV 150
Db 117 F--TGGTIDKLE----SINVTVDIDLKNSKTLNANMFVIGQTNDIVPADKLLYALRDV 170
QY 151 --TVQELDLQARHYLHGKFG 169
Db 171 TCTVDSLPLIARASILSKFAL 191

RESULT 34
A53310
pheromone CAD1 binding protein precursor - Enterococcus faecalis plasmid pAD1
N:Alternate names: TraC
C:Species: Enterococcus faecalis
C:Date: 08-Sep-1995 #sequence_revision 08-Sep-1995 #text_change 20-Aug-1999
C:Accession: A53310
R:Tanimoto, K.; An, F.Y.; Clewell, D.B.
J. Bacteriol. 175, 5260-5264, 1993
A:Title: Characterization of the traC determinant of the Enterococcus faecalis hemolysin
A:Reference number: A53310; MUID:93352433; PMID:8349566
A:Accession: A53310
A:Status: preliminary
A:Molecule type: DNA

A:Residues: 1-543 <TAN>
A:Cross-references: GB:L19532; NID:g388267; PIDN:AAA72024.1; PID:g388269
C:Genetics:
A:Gene: traC
A:Genome: plasmid
C:Superfamily: dipeptide transport protein

Query Match 7.2%; Score 88; DB 2; Length 543;
Best Local Similarity 21.3%; Pred. No. 12;
Matches 46; Conservative 31; Mismatches 61; Indels 78; Gaps 10;

QY 33 NSKAITSEKSADQFLTNTLLFKGFTTGH-----PWYNDLLVDLGGSTAATSEY- 80
Db 146 NAKETASQKSQE-----TLAVKS--NGNKTIEIELEKPTVPYFTDLL-----AUTAYVP 192
QY 81 -EGSSVDLYGAYYQCAGGTPNKTCACMYGGVTLHDNNRLTEKKVPINLWIDGKQTTVP 139
Db 193 VQOKAIKEYGKYD-----GTSQKSVITNGAFNLSLEGVGTSDK-----W-----TIS 235
QY 140 IDKVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGGKVGQRLIVFHSSEGSTVSYDLF 199
Db 236 KNKEYWDQREVSMEKINFQVVKETNGINLYN----- 267

QY 200 DAQGOYPTD-----LLRIYRDNTTISSTSLSLYL 230
Db 268 --DGOLDPTVTGEYAKQYKONKEFSTTLTMANTMYL 301

RESULT 35
A71076
hypothetical protein PH0873 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 05-Nov-1999
C:Accession: A71076
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Se
M.; Ohuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Ogu
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophil
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: A71076
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-788 <KAW>
A:Cross-references: GB:AP000004; NID:g3236131; PIDN:BAA29967.1; PID:d1030910; PID:g32
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBa
C:Genetics:
A:Gene: PH0873

Query Match 7.2%; Score 88; DB 2; Length 788;
Best Local Similarity 23.3%; Pred. No. 19;
Matches 48; Conservative 36; Mismatches 66; Indels 56; Gaps 10;

QY 5 BEINEKDLRKSELOCTALGNLQIYYNSKAITSEKSADQFLNTLLFKGFTTGHWPY 64
Db 180 EELNEENFRKELEEG--IKTLLEVAYYLKVI---DKENEYEINNLKPK----- 225
QY 65 NDLLVDLGSTAATSEYEGSSVDLYGAYYQCAGGTPNKTCACMYGGVTLHDNNRLTEKK 124
Db 226 NELKNDLSYLEASGYKYSQSQISOIFYVALKTG-----FVILSGLTSGSKTKIAQEL- 278
QY 125 VPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQA-----RH-----YLHGK----- 166
Db 279 -----GELDLDP--QVMSASGNTVAKREIKSLQETINRHGFVAVYGGWHPGKISKIK 328
QY 167 ----FGLYNS----DSFGKVGQRLIV 185
Db 329 PPFIWFVWDSDENDYKKVPYGIIV 354

RESULT 36
AE1449
hypothetical protein lin0132 [imported] - Listeria innocua (strain Clip11262)

C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE1449
R:Glaser, P.; Prangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, O.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreit, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma, O.K.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AE1449
A>Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-1946 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC95365.1; PID:g16412551; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0132

Query Match 7.2%; Score 87.5; DB 2; Length 1946;
Best Local Similarity 22.0%; Pred. No. 69;
Matches 56; Conservative 30; Mismatches 98; Indels 71; Gaps 12;

QY 4 SEEINEKDLRKSELOGTALGNLKOIYY-----YNSKAITSSEKSADQFLTNTLL 53
DB 1198 TRLNLRADYRVTKGPTDSEGNAGVKFKNIWTKYIAYRGGMISNSETADPF-----S 1252

QY 54 FK----GFTGHPWYNDLLVLDGSAATSEYEGSSVDLYGAYGYQCAGGT-----P 101
DB 1253 FRQEDGSYVAB--UNGRYVDFSSGLVSTEAGSKIDLYSAQEGVSQVSAITTKALVEP 1310

QY 102 NKTACMYGVTLLHNNRLTEEEKVPINLWIDGKQTPVPIDKVKTSKKEVTQO---ELDIQ 158
DB 1311 NITA---EVSIVDRR-----VKLDVVGEDKNTTVKKDNENELF 1347

QY 159 ARHYLHGFGLYNSDFSGKQVORGLIVPHSS-----EGSTVSYDLFDAQGYPDTL 209
DB 1348 VNAYREDGETLVKSVRIDGLPTDFIVTDLSPDKYVIRKVEG---KYDLLDGEQ--PQN- 1401

QY 210 LRIYRDNNTTISSTSL 224
DB 1402 -KVVYSETITTEKSL 1415

RESULT 37
T08081
probable myrosinase-binding protein - rape
A:Alternate names: jasmonate inducible protein
C:Species: *Brassica napus* (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 11-May-2000
C:Accession: T08081
R:Geshi, N.; Brandt, A.
Planta 204, 295-304, 1998
A:Title: Two jasmonate-inducible myrosinase-binding proteins from *Brassica napus* L. seed
A:Reference number: Z16340; MUID:98192006; PMID:9530873
A:Accession: T08081
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-914 <GES>
A:Cross-references: EMBL:Y11482; NID:el023100; PIDN:CAA72270.1; PID:e304149
A:Experimental source: cv. Global; isolate a4; seedlings
A>Note: jasmonate inducible

Query Match 7.1%; Score 87; DB 2; Length 914;
Best Local Similarity 20.1%; Pred. No. 28;
Matches 65; Conservative 35; Mismatches 89; Indels 134; Gaps 14;

QY 2 EKSEINE-----KDLRKSELOGTALGNLKOIYY-YNS 34
DB 347 EKGEFNDFGEGVKKITVGADQYSVTYKIEYKDGKVEIREHGTSGELQEFSDVDPN 406

QY 35 KAITSSEKSADQFLT--NTLLFKGFTGHPWYNDLLVLDGSAATSEYE----- 81

Db 407 DSITEVGGTYKHNYYDTTLTSLYFTTSKGFTSPLFGINSEKKGTFFEFKDENGGKLIG 466
QY 82 -----GSSVDLYGAYGYQCAGG-----TPNKTACMYGGVTLHNNRLTEE 122
DB 467 LHGRGNAIDALGAYFDFTGSGQGDGSDVPSKDGPTDVPGRKGPLGG-----DKGEPPDD 521

QY 123 -----KKVPI---NLWIDGKQTPVPIDKVKTSKKEV-----TVQ 153
DB 522 VGFEGVKKITVGADNLSI---TYKIEYKDGKVEIREHGTARGKIKKEFSVDYDNDST 577

QY 154 ELDLARH-----YLHGKFGLYN-----SDSFGKGVORGLIV 185
DB 578 EVGGTYKHNYYDTTLTSLYFTTSKGFTSPLFGIDSEKKGTFFEFKDENGGK----LIG 633

QY 186 FHSSEGSTVSYDLFDAQGYPD 208
DB 634 FHGRGNAI-----DAIGAYFDT 651

RESULT 38
T08102
myrosinase-binding protein (clone indmbp2) - rape
C:Species: *Brassica napus* (rape)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 21-Jul-2000
C:Accession: T08102
R:Taipalensuu, J.; Eriksson, S.J.; Rask, L.
Eur. J. Biochem. 250, 680-688, 1997
A:Title: The myrosinase binding protein from *Brassica napus* seeds possesses lectin ac
A:Reference number: Z16355; MUID:98121188; PMID:9461290
A:Accession: T08102
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-988 <TAI>
A:Cross-references: EMBL:Y09437; NID:gl711295; PIDN:CAA70587.1; PID:gl711296
A:Experimental source: cv. Hanna; leaves

Query Match 7.1%; Score 87; DB 2; Length 988;
Best Local Similarity 20.5%; Pred. No. 31;
Matches 66; Conservative 35; Mismatches 89; Indels 132; Gaps 15;

QY 2 EKSEINE-----KDLRKSELOGTALGNLKOIYY-YNS 34
DB 363 EKGEFNDFGEGVKKITVGADQYSVTYKIEYKDGKVEIREHGTSGELQEFSDVDPN 422

QY 35 KAITSSEKSADQFLT--NTLLFKGFTGHPWYNDLLVLDGSAATSEYE----- 81
DB 423 DSITEVGGTYKHNYYDTTLTSLYFTTSKGFTSPLFGINSEKKGTFFEFKDENGGKLIG 482

QY 82 -----GSSVDLYGAYGYQCAGG-----TPNKTACM-----YGGVTLHNNRLTEE- 122
DB 483 LHGRGNAIDALGAYFDFTGSGQGDGSDVPTKDVLIADVPGRKGPLGG-----DKGEPPDDV 538

QY 123 -----KKVPI---NLWIDGKQTPVPIDKVKTSKKEV-----TVQE 154
DB 539 GFEGVKKITVGADNLSI---TYKIEYKDGKVEIREHGTARGKIKKEFSVDYDNDST 594

QY 155 LDLARH-----YLHGKFGLYN-----SDSFGKGVORGLIV 186
DB 595 VGGTYKHNYYDTTLTSLYFTTSKGFTSPLFGIDSEKKGTFFEFKDENGGK----LIG 650

QY 187 HSSEGSTVSYDLFDAQGYPD 208
DB 651 HGRGNAI-----DAIGAYFDT 667

RESULT 39
E71403
hypothetical protein - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cross)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: E71403

R;Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk P.; Wedler, H.; Wedler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terryn, N.; Gielavanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Rieger, M.; Schaeffer, M.; Funk, B. Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech erhoft, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans C.; Chalwatzis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thaliana

A:Reference number: A71400; MUID:98121113; PMID:9461215

A:Accession: E71403

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-722 <BEV>

A:Cross-references: GB:297335; NID:g2244747; PID:e326875; PID:g2244776

C:Genetics:

A:Map position: 4COP9-4C3845

Query Match 7.1%; Score 86.5; DB 2; Length 722;
Best Local Similarity 20.2%; Pred. No. 23;
Matches 47; Conservative 40; Mismatches 79; Indels 67; Gaps 10;

QY 6 EINEKDLRKSELOGTAGLNKUKQIYYNNAKITSSEKSADQFLNTLLFKGFFTGHPWYN 65
Db 455 KINSEDIGMKAEEGKSLGNAQTV---ESETLRVSLPAVD----- 491

QY 66 DLLVDLGSTAAATSEYSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNR---LTE 121
Db 492 SVVVDNADSSAANKTGLVDLAG-----NESELQANVAEKGKPKDILLSS 539

QY 122 EKKV---PINLWIDGKQTTVPIDKVKTSKKEVTQVELDQARHYLHGKFGLYNSDSFGGK 178
Db 540 ESSCYASPVSVFEGSDAS---DOIKSSTSKDSALQISAE-----SKDEVCRE 586

QY 179 VORGLIVPHSS---EGSTVSYDL--FDAQGYPTLRIYRDNTTISTSLSI 226
Db 587 SNNGRLEVESSFINEPNKTEYPISFGSTGTAPDTV-----NTANQTSLEI 632

RESULT 40
T30857
glucosyltransferase - Streptococcus salivarius
C:Species: Streptococcus salivarius
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30857
R:Simpson, C.L.; Giffard, P.M.; Jacques, N.A. Infect. Immun. 63, 609-621, 1995

A:Title: Streptococcus salivarius ATCC 25975 possesses at least two genes coding for proteinase

A:Reference number: Z20909; MUID:95122197; PMID:7822030

A:Accession: T30857

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1449 <SIM>

A:Cross-references: EMBL:L35495; NID:g662378; PID:g662379; PIDN:AAC41412.1

C:Genetics:

A:Gene: gtfL

Query Match 7.1%; Score 86.5; DB 2; Length 1449;
Best Local Similarity 25.0%; Pred. No. 57;
Matches 46; Conservative 26; Mismatches 75; Indels 37; Gaps 9;

QY 23 LGNLKQIYYNNAKITSSEKSADQFLNTLL-----LFKGFFTGHPWY- 64
Db 245 LSNQKE-YRFDKNGGTGS--SADSTNTNTVNGDKNAFYGTTDKDIELVDGYFTANTWYR 301

QY 65 -NOLLVDLGSTAAATSEY-----GSSVDLYGAYGYQCAGGT-PNKTACMYGGVTLH 114
Db 302 PKEILKDGKEWTASTENDKRPLLTVWNPWSKAIOASYLNMYNKEQGLGTNQTVTSFSSQTOM 361

QY 115 DNNRLTEKKVPINLWIDGK---QTTVPIDKVKTSKKEVTQVELDQARHYLHGKFGLY 170
Db 362 DQAALYVOKRIEERIEAREGNTDWLRTTIK-NFVKTPQGNWSTSE-NLDNNDHLOGGALLY 419

QY 171 NSDS 174

Db 420 NDS 423

RESULT 41
T38720
chromodomain helicase - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000

C:Accession: T38720

R:Gentiles, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V. submitted to the EMBL Data Library, September 1997

A:Reference number: Z21797

A:Accession: T38720

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1388 <GEN>

A:Cross-references: EMBL:Z99167; PIDN:CAB16277.1; GSPDB:GN00066; SPDB:SPAC3G6.01

A:Experimental source: strain 972h-; cosmid c3G6

C:Genetics:

A:Gene: SPDB:SPAC3G6.01

A:Map position: 1

A:Introns: 64/3

C:Superfamily: CHD-1 protein; chromobox homology

Query Match 7.1%; Score 86; DB 2; Length 1388;
Best Local Similarity 25.1%; Pred. No. 60;
Matches 45; Conservative 27; Mismatches 69; Indels 38; Gaps 10;

QY 35 KAITSSSEKSADQFLTNT-----LLPKGF-----FTGHPWYNDLLVDLGSTAAATSEVE 81
Db 1047 KAVSAAEKDLSNDOSNNKSSRKALLITFKGVKNNAETLVORLNDLDI-LYDAMPTSGYS 1105

QY 82 GSSVDLY-GAYGYQCAGGTPNKTACMYGGVTLH-----DNNRLTEKKVPINLWI 131
Db 1106 NFQIPMHVRSVHGWSQWG-PREDSMLLSGICKHGFGAWLEIRDDPELKMKDIFLE--- 1161

QY 132 DGKOT--TVPIDKVKTSKKEVTQVELDQARHYL-----HGKFGLYNSD--SFGGKQV 180
Db 1162 DTQOTDNSVPKDKENKEKKVPKSAVHLVRGEYLLSALREHQNFQIKSPAISTNGKTQ 1220

RESULT 42
AH0340
putative autotransporter protein yapC [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001

C:Accession: AH0340

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M. deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G. il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrel Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AH0340

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-638 <KUR>

A:Cross-references: GB:AL50842; PIDN:CAC93031.1; PID:g15980769; GSPDB:GN00175

C:Genetics:

A:Gene: yapC

Query Match 7.0%; Score 85.5; DB 2; Length 638;
Best Local Similarity 21.1%; Pred. No. 24;
Matches 43; Conservative 20; Mismatches 70; Indels 71; Gaps 6;

QY 68 LVDLG--STAATSEYSSVDLYGAYGYQCAGGTPNKTACMYGGVT----- 112
Db 299 VVDLGVQYGLYSQESNGSTDWYLATSTTELPFGTTPNVTAPMLSSAAOGLVNLMAAPRHI 358

QY 113 -----LHDNNRLTEKKVPINLWIDGKQTTVPIDKV 143
Db 359 LNAELSTLRQGEKADAEAGTGVGWARYLTDSDRLSDNKNIAFKNTLSGME--IGADK- 415

QY 144 KTSKKEVTVOELDLQARHYLHGKFLGYNDSFGKVGORGLIVFHSSEGSTVS- D 197
Matches 64; Conservative 35; Mismatches 100; Indels 98; Gaps 15;
QY 416 -----QGLNNGNMLGIFTSSSDVKST-----HDANGDIRSYGGGLYLT 457
QY 198 LFDAGQGYPTDLLRIYRONTTISS 221
Db 458 YLQSGFYVDTVLKANRFNKNMT 481
RESULT 43
T30552
glucosyltransferase N - Streptococcus salivarius (fragment)
C:Species: Streptococcus salivarius
C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: T30552
R:Jaffe, R.I.
submitted to the EMBL Data Library, February 1998
A:Description: Streptococcus salivarius V1477 gtfN.
A:Reference number: Z20854
A:Accession: T30552
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1449 <JAF>
A:Cross-References: EMBL:AF049609; NID:g2935545; PID:g2935546; PIDN:AAC05156.1
C:Genetics:
A:Gene: gtfN
Query Match 7.0%; Score 85.5; DB 2; Length 1449;
Best Local Similarity 25.0%; Pred. No. 70;
Matches 46; Conservative 26; Mismatches 75; Indels 37; Gaps 9;
QY 23 LGNLKOIYYNKAITSSEKSAQDLTNL-----LFGKFTGHPW- 64
Db 245 LSNQKE-YFDXNGGTGS--SADSTNTVTVNGDKNAFTGTDKDIELVDGFTTATWTR 301
QY 65 -NDLLVDLGLSTAATSEYE-----GSSVDLYGAYGYQCAGGT-PNKTACMYGGVTLH 114
Db 302 PREILKDGREWTASTENDKRPLLVNWPMSKAIQASLYNWKKEGLTNGTYTSFSSQTQM 361
QY 115 DNNRLTEKKVPINLWIDCK-----QTTVPIDKVKTSKEVTVOELDLQARHYLHGKFLG 170
Db 362 DQAALVEQKRIEGRIAREGNTOWLRTTIK-NFVKTPGWNSTSE-NLDNNDHLQGGALLY 419
QY 171 NSDS 174
Db 420 NND5 423
RESULT 44
AG1749
glycosidase homolog lin2540 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C>Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG1749
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Anand, A.; Baquero, F.; Berche, P.; Bloeker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1749
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1090 <GLA>
A:Cross-References: EMBL:AL592022; PIDN:CAC97767.1; PID:gl61415062; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2540
Query Match 7.0%; Score 85; DB 2; Length 1090;

Best Local Similarity 21.5%; Pred. No. 53;
Matches 64; Conservative 35; Mismatches 100; Indels 98; Gaps 15;
QY 8 NEKDLRKKSELOGTALGNLK-----QIYYNKAITSSEKSAQDLTN 50
Db 400 NEWD--RESDV-STALSNAKENEIPATGVLEQWSEDETYIWNDAITYAKKNGEFPAYD 456
QY 51 TLLFKGFFTGHPWYNDLLVDLGLSTAA-----TSEYEGSSVDL-YGAYGYQCAGG 99
Db 457 DTFNGKWTNPKAMWDSVHDAGMNVILWQVPLVKDDNTVVEQRDDEDEYAIISGYASDDG 516
QY 100 T-PNKTACMY---GGVTLHDNNRLTEKKVPINLWIDCKQ---TTVPIDKVKTSKEVT 151
Db 517 TGAHYRVPASQWFGNGILLDFTN-----KDAVNNWTSQREYLLTDVGIDGFKTDGGMV 570
QY 152 -----VQELDLQARHYLHG-----KFLGYN-- 172
Db 571 WGRDTTFSNGEKQEMNRNRYPTDYVSSYDFFAKSIINFAVSEFSRSGTSQAQSGIYWSGD 630
QY 173 -----DSFGKVGORGLIVFHSSEGSTVS- -DLFDAQGOYPTDLLRIYRDNTTIS 221
Db 631 QTSTFDSFOASVKAGL-----SASTSGVSWANDMAGFTGNYPTA--ELYKRATAMAA 681
RESULT 45
S23771
outer membrane porin protein OprD precursor PA0958 [imported] - Pseudomonas aeruginosa
N:Alternate names: pore-forming protein; protein D2
C:Species: Pseudomonas aeruginosa
C>Date: 22-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 15-Jun-2001
C:Accession: S23771; S61560; S25008; E83527
R:Yoneyama, H.; Yoshihara, E.; Nakae, T.
Antimicrob. Agents Chemother. 36, 1791-1793, 1992
A:Title: Nucleotide sequence of the protein D2 gene of Pseudomonas aeruginosa.
A:Reference number: S23771; MUID:93037310; PMID:1339257
A:Accession: S23771
A:Molecule type: DNA
A:Residues: 1-443 <YON>
A:Cross-References: EMBL:X63152; NID:g45370; PIDN:CAA44855.1; PID:g45371
A:Accession: S61560
A:Molecule type: protein
A:Residues: 24-34 <YOM>
R:Huang, H.; Siehnell, R.J.; Bellido, F.; Rawling, E.; Hancock, R.E.W.
submitted to the EMBL Data Library, July 1992
A:Description: Analysis of two gene regions involved in the expression of the Imipene
A:Reference number: S23859
A:Accession: S25008
A:Molecule type: DNA
A:Residues: 1-443 <HUA>
A:Cross-References: EMBL:Z14065; NID:g45372; PIDN:CAA78448.1; PID:g45373
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: E83527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-443 <STO>
A:Cross-References: GB:AE004529; GB:AE004091; NID:g9946851; PIDN:AAG04347.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: oprD; PA0958
C:Superfamily: Pseudomonas aeruginosa outer membrane porin PA0958
C:Keywords: porin
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-443/Product: protein D2 #status experimental <MAT>
Query Match 6.9%; Score 84.5; DB 2; Length 443;
Best Local Similarity 23.2%; Pred. No. 18;
Matches 60; Conservative 30; Mismatches 98; Indels 71; Gaps 14;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run On: January 8, 2003, 11:52:22 ; Search time 4.63425 seconds
(without alignments)
2085.338 Million cell updates/sec

Title: US-09-900-766-2

Perfect score: 1218

Sequence: 1 SEKSEINEKDLRKSELQ.....RDNTTISTSLSLYLTYT 233

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1107.	90.9	257	1 ETXE_STAAU	P12993 staphylococ
2	948	77.8	257	1 ETXA_STAAW	P13163 staphylococ
3	611	50.2	258	1 ETXD_STAAU	P20723 staphylococ
4	284.5	23.4	266	1 ETC3_STAAH	P23313 staphylococ
5	277.5	22.8	266	1 ETC2_STAAU	P34071 staphylococ
6	269.5	22.1	251	1 SPEA_STRPY	P08095 streptococ
7	268.5	22.0	266	1 ETXB_STAAU	P01552 staphylococ
8	263.5	21.6	266	1 ETC1_STAAU	P01553 staphylococ
9	241	19.8	258	1 ETXG_STAAH	O85382 staphylococ
10	185	15.2	236	1 SPEH_STRPY	Q9X5C8 streptococ
11	183.5	15.1	234	1 SPEG_STRPY	Q9X5C7 streptococ
12	148.5	12.2	235	1 SPEC_STRPY	P13380 streptococ
13	100	8.2	234	1 TSST_STAAU	P06886 staphylococ
14	95.5	7.8	150	1 R19E_THEVO	Q97CU4 thermoplas
15	93	7.6	979	1 P115_MYCHR	P41508 mycoplasma
16	91	7.5	1634	1 DPOL_METUA	Q58295 methanococ
17	88.5	7.3	150	1 R19E_THEAC	Q9Hm21 thermoplas
18	88	7.2	419	1 TVPH_MYCPT	P47717 mycoplasma
19	86	7.1	1388	1 HRP3_SCHPO	O14139 schizosach
20	84.5	6.9	443	1 PORD_PSEAE	P32722 pseudomonas
21	84	6.9	590	1 DNKX_STRMU	O06942 streptococ
22	84	6.9	1050	1 ITA5_XENLA	Q06274 xenopus lae
23	83.5	6.9	282	1 PRTA_ASPNG	P24665 aspergillus
24	82.5	6.8	328	1 YB70_HAEIN	Q57527 haemophilus
25	82.5	6.8	455	1 YWAD_BACSU	P25152 bacillus su
26	82.5	6.8	600	1 NUCD_ECOLI	P33599 escherichia
27	82	6.7	607	1 DNKX_STRPY	P95831 streptococ
28	82	6.7	843	1 PULA_THEMA	O33840 thermotoga
29	82	6.7	3718	1 LMA5_MOUSE	Q61001 mus musculu
30	81	6.7	609	1 DNKX_STRAG	P95693 streptococ
31	81	6.7	611	1 WDR1_CAEEL	Q11176 caenorhabdi
32	80	6.6	345	1 PURA_METUA	Q57981 methanococ
33	80	6.6	411	1 PGK_METTH	O27121 methanobact

ALIGNMENTS

RESULT 1

```

ETXE_STAAU
ID ETXE_STAAU STANDARD; PRT; 257 AA.
AC P12993;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type E precursor (SEE).
GN ENTE.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcu.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-74.
RC STRAIN=MJR265;
RX MEDLINE=88257005; PubMed=3384800;
RA Couch J.L., Soltis M.T., Betley M.J.;
RT "Cloning and nucleotide sequence of the type E staphylococcal
enterotoxin gene.";
RL J. Bacteriol. 170:2954-2960(1988).
RN [2]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=96022387; PubMed=7552730;
Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
"Residues defining V beta specificity in staphylococcal
enterotoxins.";
RL Nat. Struct. Biol. 2:680-686(1995).
CC -!- CORACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
for the toxin interaction with MHC class II (By similarity).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M21319; AAA26617.1; -.
PIR; A28179; A28179.
PDB; 1SEE; 15-OCT-95.
InterPro; IPR001961; Staph/Strep_toxin.
Pfam; PF01123; Staph_Strep_toxin; 1.
Pfam; PF02876; Staph_Strep_tox_C; 1.
PRINTS; PR00279; BACTRLTOXIN.
PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
Enterotoxin; Toxin; Signal; Supertoxin; Metal-binding; zinc;
3D-structure; 1 27
SIGNAL CHAIN 28 257 ENTEROTOXIN TYPE E.

```

```
FT METAL 211 211 ZINC (BY SIMILARITY).
FT METAL 249 249 ZINC (BY SIMILARITY).
FT METAL 251 251 ZINC (BY SIMILARITY).
SQ SEQUENCE 257 AA; 29358 MW; 27EDA94B97770CE3 CRC64;

Query Match 90.9%; Score 1107; DB 1; Length 257;
Best Local Similarity 89.7%; Pred. No. 2.2e-90;
Matches 209; Conservative 9; Mismatches 15; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKKSELOQTALGNLKIYYNNKAITSSSEKSADQFTNTLLFKGFTTG 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 SEKSEINEKDLRKKSELOQTALGNLKIYYNNKAITENKESDDQFLENTLLFKGFTTG 84
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 85 HPWYNDLLVDLGSKDATNKYKGGKVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 144
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 121 ERKKVPINLWDGKOTTPIDKVKTSKKEVTVOELDLQARHVLHGKFGLYNSDSFGKQV 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 145 ERKKVPINLWDGKOTTPIDKVKTSKKEVTVOELDLQARHVLHGKFGLYNSDSFGKQV 204
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 181 RGLIVFHSSEGSTVSYDLFDAGQYPTDLLRIYRDNKTISSTLSISLYLYTT 233
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 205 RGLIVFHSSEGSTVSYDLFDAGQYPTDLLRIYRDNKTINSENHIDLXYLYTT 257
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 2
ID ETVA_STAAW STANDARD; PRT; 257 AA.
AC P13163;
DC 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type A precursor (SEA).
GN ENTA OR MW1889.
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=196620, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwama N., Asano K., Naimi T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RL Lancet 359:1819-1827(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=FR1337;
RX MEDLINE=88086892; PubMed=3335483;
RA Betley M.J., Mekalanos J.J.;
RT "Nucleotide sequence of the type A staphylococcal enterotoxin gene.";
RL J. Bacteriol. 170:34-41(1988).
RN [3]
RP SEQUENCE OF 25-257.
RX MEDLINE=87222293; PubMed=3584106;
RA Huang I.-Y., Hughes J.L., Bergdoll M.S., Schantz E.J.;
RT "Complete amino acid sequence of staphylococcal enterotoxin A.";
RL J. Biol. Chem. 262:7006-7013(1987).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=95354648; PubMed=7628431;
RA Schad E.M., Zaitseva I., Zaitsev V.N., Dohlisten M., Kalland T.,
RA Schlievert P.M., Ohlendorf D.H., Svensson L.A.;
RT "Crystal structure of the superantigen staphylococcal enterotoxin
RT type A.";
RL EMBO J. 14:3292-3301(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).
RX MEDLINE=97113025; PubMed=8943278;
```

```
RA Sundstroem M., Hallen D., Svensson A., Schad E., Dohlisten M.,
RA Abrahamson L.;
RT "The Co-crystal structure of staphylococcal enterotoxin type A with
RT Zn2+ at 2.7-A resolution. Implications for major histocompatibility
RT complex class II binding.";
RL J. Biol. Chem. 271:32212-32216(1996).
RN [6]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=96022987; PubMed=7552730;
RA Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
RT "Residues defining V beta specificity in staphylococcal
RT enterotoxins.";
RL Nat. Struct. Biol. 2:680-686(1995).
RN [7]
RP COMPARISON OF STRUCTURE OF SEA AND SEC2.
RX MEDLINE=97334373; PubMed=9191070;
RA Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;
RT "A structural and functional comparison of staphylococcal
RT enterotoxins A and C2 reveals remarkable similarity and
RT dissimilarity.";
RL J. Mol. Biol. 269:270-280(1997).
CC -!- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary
CC for the toxin interaction with MHC class II.
CC -!- SUBUNIT: MONOMER.
CC -!- SUBCELLULAR LOCATION: Secreted
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY A BACTERIOPHAGE.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AP004828; BAB95754.1; -.
DR EMBL; M18970; AAA26681.1; -.
DR PIR; A28664; A28664.
DR PIR; A29566; A29566.
DR PDB; 1ESF; 1I-JUL-96.
DR PDB; 1SEA; 15-OCT-95.
DR PDB; 1SXT; 19-NOV-97.
DR InterPro; IPR001961; Stap/Strep_toxin.
DR Pfam; PF01123; Stap_Strep_toxin; 1.
DR Pfam; PF02876; Stap_Strep_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;
KW 3D-structure.
FT SIGNAL 1 24
FT CHAIN 25 257 ENTEROTOXIN TYPE A.
FT DISULFID 120 130
FT METAL 211 211 ZINC.
FT METAL 249 249 ZINC.
FT METAL 251 251 ZINC.
FT CONFLICT 242 242 T -> S (IN REF. 3).
SQ SEQUENCE 257 AA; 29669 MW; ADEBF5BCALF14677 CRC64;

Query Match 77.8%; Score 948; DB 1; Length 257;
Best Local Similarity 76.4%; Pred. No. 2.3e-76;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

Qy 1 SEKSEINEKDLRKKSELOQTALGNLKIYYNNKAITSSSEKSADQFTNTLLFKGFTTG 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 25 SEKSEINEKDLRKKSELOQTALGNLKIYYNNKAITENKESDDQFLENTLLFKGFTTG 84
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLT 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Db 85 HSWYNDLVDFDSKIDVVKYKGGKVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 144
QY 121 EEKVPINLWDGKQTPIDKVKTSKKEVTVOELDLQARHLYHGKFLGYNLSDSFGGKVQ 180
Db 145 EEKVPINLWDGKQTPLETVTKNTKNVTVOELDLQARHLYHGKFLGYNLSDSFGGKVQ 204
QY 181 RGLIVFHSSEGSTVSYDLFDAGQGVPTDLLRIYRDNNTTISSTLSISLYLT 233
Db 205 RGLIVFHTSTEPSVNYDLFGAGQGVNTLLRIYRDNNTTINSENMHIDIVLTS 257

RESULT 3

ETXD_STAAM STANDARD: PRT; 258 AA.
AC P20723;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type D precursor (SED).
GN ENT.D.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89359112; PubMed=2549000;
RA Bayles K.W., Iandolo J.J.;
RT "Genetic and molecular analyses of the gene encoding staphylococcal enterotoxin D";
RL J. Bacteriol. 171:4799-4806(1989).
RN [2]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX STRAIN=ATCC 23235;
RX MEDLINE=97157473; PubMed=9003758;
RA Sundstroem M., Abrahamsen L., Antonsson P., Mehindate K., Mourad W.,
RA Dohlsten M.;
RT "The crystal structure of staphylococcal enterotoxin type D reveals Zn²⁺-mediated homodimerization";
RL EMBO J. 15:6832-6840(1996).
CC -!- SUBUNIT: HOMODIMER; ZINC-DEPENDENT.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M28521; AAB06195.1; -;
DR PIR; A33953; A33953.
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph/Strep toxin.
DR Pfam; PF01123; Staph_Strp_toxin; 1.
DR Pfam; PF02876; Staph_Strp_cox_C; 1.
DR PRINTS; PR00279; BACTRIP_TOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superantigen; Zinc.
FT SIGNAL 1 25
FT CHAIN 26 258 ENTEROTOXIN TYPE D.
FT METAL 212 212 ZINC.
FT METAL 250 250 ZINC.
FT METAL 252 252 ZINC.
FT VARIANT 114 114 P -> A (IN STRAIN ATCC 23235).
SQ SEQUENCE 258 AA; 29746 MW; 4F7C6A28D42597FD CRC64;

Query Match

50.2%; Score 611; DB 1; Length 258;

Best Local Similarity 51.1%; Pred. No. 1.le-46;
Matches 118; Conservative 35; Mismatches 78; Indels 0; Gaps 0;
QY 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNSKAITSSSEKSAQOFLTNLLFKGFFTG 60
Db 26 NENIDSVKKEKHLKKSSELSSTALNNKHSYADKNPIIGENKSTGDFLENTLLYKFFTD 85
QY 61 HPWYNDLVLDLSTAAATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLT 120
Db 86 LINFEDLLINFSNKEMAQHFKSKNVDPYPIRISINCYGGEIDRTACTYGCVTPHGNK 145
QY 121 EEKVPINLWDGKQTPIDKVKTSKKEVTVOELDLQARHLYHGKFLGYNLSDSFGGKVQ 180
Db 146 ERKKIPINLWINGVQKVS LDKVQTDKKNVTVOELDLQARHLYHGKFLGYNLSDSFGGKVQ 205
QY 181 RGLIVFHSSEGSTVSYDLFDAGQGVPTDLLRIYRDNNTTISSTLSISLYLY 231
Db 206 RGLIEFSDSGSKVSYDLFDVKGDFPEKQLRIYSDNKTLSLSTHLLHIDIVLY 256

RESULT 4

ETC3_STAAM STANDARD: PRT; 266 AA.
ID ETC3_STAAM
AC P23313;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type C-3 precursor (SEC3).
GN ETC3 OR SAV2009 OR SA1817.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Murayama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus aureus";
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220508; PubMed=2325627;
RA Hovde C.J., Hackett S.P., Bohach G.A.;
RT "Nucleotide sequence of the staphylococcal enterotoxin C3 gene: sequence comparison of all three type C staphylococcal enterotoxins";
RL Mol. Gen. Genet. 220:329-333(1990).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF COMPLEX WITH TCR.
RX MEDLINE=97064178; PubMed=8906797;
RA Fields B.A., Malchiodi E.L., Li H., Ysern X., Stauffacher C.V.,
RA Schlievert P.M., Karjalainen K., Mariuzza R.A.;
RT "Crystal structure of a T-cell receptor beta-chain complexed with a superantigen";
RL Nature 384:188-192(1996).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; AP003364; BAB58171.1; -
DR EMBL; AP003135; BAB43097.1; -
DR EMBL; X51661; CAA35972.1; -
DR PIR; S11885; S11885.
DR PDB; 1JCK; 12-NOV-97.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Stap_Strp_toxin; 1.
DR Pfam; PF02876; Stap_Strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superantigen; 3D-structure;
KW Complete proteome.
FT SIGNAL 1 27
FT CHAIN 28 266 ENTEROTOXIN TYPE C-3.
FT DISULFID 120 137
SQ SEQUENCE 266 AA; 30671 MW; 5ED8A32D11FCA59 CRC64;

Query Match 23.4%; Score 284.5; DB 1; Length 266;
Best Local Similarity 32.5%; Pred. No. 6.5e-18;
Matches 76; Conservative 45; Mismatches 94; Indels 19; Gaps 9;

QY 11 DLKKSELOQTALGNLKOIYYNSKATSSP-KSADQFLNTLLFKGFTTGHVYNDLLV 69
DB 37 DLKSSSEFTGT-MGNMK--YLYDDHYVSATKVKSVDFLAHDLIYNISDKKLKKNYDKVKT 93
QY 70 DLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTAGMGVGTLDHNNRLTEE 122
DB 94 ELLNEDLAKKYKDEVVDVYGSNYVNCYFSSKDNVGVGTGKTCMYGGITKHEGNHFDNG 153
QY 123 --KKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLYNSDFSFGKVKV 180
DB 154 NLQNVLRVY-ENKRNTISFE-VQTDKKSVAQELDIKARNFLINKKNLYEENS--SPYE 209
QY 181 RGLVHFSSEGSTVSYDLFDAQGOYPD--TLRLRYRDNNTTSSLSLSLYLT 232
DB 210 TGYIKFIENNGNTFWYDMMPAPGDKFQOSKYLMMYNDNKNKTVDKSVKIEVHLT 263

RESULT 5
ETC2_STAAU STANDARD; PRT; 266 AA.
AC P34071;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Enterotoxin type C-2 precursor (SEC2).
GN ENT2C.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-66.
RX MEDLINE=96027099; PubMed=2543637;
RA Bohach G.A., Schlievert P.M.;
RT "Conservation of the biologically active portions of staphylococcal enterotoxins C1 and C2.";
RL Infect. Immun. 57:2249-2252(1989).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
RX MEDLINE=96027099; PubMed=7582894;
RA Papageorgiou A.C., Acharya K.R., Shapiro R., Passalacqua E.F.,
RA Brehm R.D., Tranter H.S.;
RT "Crystal structure of the superantigen enterotoxin C2 from Staphylococcus aureus reveals a zinc-binding site.";
RL Structure 3:769-779(1995).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).

RX MEDLINE=96022987; PubMed=7552730;
RA Swaminathan S., Furey W.F. Jr., Pletcher J., Sax M.;
RT "Residues defining V beta specificity in staphylococcal enterotoxins.";
RL Nat. Struct. Biol. 2:680-686(1995).
RN [4]
RP COMPARISON OF STRUCTURE OF SEA AND SEC2.
RX MEDLINE=97334373; PubMed=9191070;
RA Schad E.M., Papageorgiou A.C., Svensson L.A., Acharya K.R.;
RT "A structural and functional comparison of staphylococcal enterotoxins A and C2 reveals remarkable similarity and dissimilarity.";
RL J. Mol. Biol. 269:270-280(1997).
CC -1- COFACTOR: Binds 1 zinc ion per subunit. The zinc ion is necessary for the toxin interaction with MHC class II.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -1- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN FAMILY.
DR PIR; A60114; A60114.
DR PDB; 1STE; 23-DEC-96.
DR PDB; 1SE2; 08-MAR-96.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Stap_Strp_toxin; 1.
DR Pfam; PF02876; Stap_Strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Signal; Superantigen; Metal-binding; Zinc;
KW 3D-structure.
FT SIGNAL 1 27 ENTEROTOXIN TYPE C-2.
FT CHAIN 28 266
FT DISULFID 120 137
FT METAL 36 36 ZINC.
FT METAL 110 110 ZINC.
FT METAL 145 145 ZINC.
FT METAL 149 149 ZINC.
SQ SEQUENCE 266 AA; 30604 MW; 8407FB18536FAC08 CRC64;
Query Match 22.8%; Score 277.5; DB 1; Length 266;
Best Local Similarity 31.1%; Pred. No. 2.7e-17;
Matches 76; Conservative 47; Mismatches 102; Indels 19; Gaps 9;
QY 1 SEKSEINEKDLRKKSLOQTALGNLKOIYYNSKATSSP-KSADQFLNTLLFKGFTT 59
DB 27 AESQPDPTDELKHSSEFTGT-MGNMK--YLYDDHYVSATKVKSVDFLAHDLIYNISDK 83
QY 60 GHPWYNDLLVDLSTAAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTAGMGVGT 112
DB 84 KLKNDYKVKTELLNEDLAKKYKDEVVDVYGSNYVNCYFSSKDNVGVGTGKTCMYGGIT 143
QY 113 LHDNNRLTEE--KKVPINLWIDGKQTTVPIDKVKTSKEVTVQELDLQARHYLHGKFGLY 170.
DB 144 KHGHNHFDNGNLQNVLRVY-ENKRNTISFE-VQTDKKSVAQELDIKARNFLINKKNLY 201
QY 171 NSDSFGGKQVQRLVHFSSEGSTVSYDLFDAQGOYPD--TLRLRYRDNNTTSSLSLSL 228
DB 202 EFN5--SPYETGYIKFIENNGNTFWYDMMPAPGDKFQOSKYLMMYNDNKNKTVDKSVKIEV 259
QY 229 YLYT 232
DB 260 HLT 263
RESULT 6
SPEA_STRPY STANDARD; PRT; 251 AA.
ID SPEA_STRPY
AC P08095;
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Exotoxin type A precursor (Scarlet fever toxin) (Erythrogenic toxin)


```
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -I- SUBUNIT: Binds to major histocompatibility complex class II beta
CC chain.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- DISEASE: Mitogenic for human peripheral blood lymphocytes.
CC -I- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF124500; AAD30989.1; -.
CC EMBL: AE006546; AAK33907.1; -.
CC HSSP: P13163; 1SXT.
CC InterPro: IPR001961; Staph/Strep_toxin.
CC Pfam: PF01123; Staph_Strp_toxin; 1.
CC Pfam: PF02876; Staph_Strp_tox_C; 1.
CC PROSITE: PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
CC PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
CC Toxin; Signal; Complete proteome.
CC SIGNAL 1 32 POTENTIAL.
CC CHAIN 33 236 EXOTOXIN TYPE H.
CC SEQUENCE 236 AA; 27485 MW; 16352923907AD40D CRC64;

Query Match 15.28; Score 185; DB 1; Length 236;
Best Local Similarity 26.78; Pred. No. 3.3e-09;
Matches 56; Conservative 46; Mismatches 84; Indels 24; Gaps 9;

Qy 25 NLKQIYYNSKAI-TSEKSDAQFLNTLLFKGFTTGHWPYN-----DLLVDLGSTAAATSE 79
Db 42 NLESYLKHDLSLEADSIKSPDVTSHML-----KYSVKDKNLVSFFERKDWISQE 92

Qy 80 YEGSSVDLYGAYGVQCAGGTPNKACMGVTLHDNNRLTEKKVPTNLWIDGKQTTPV 139
Db 93 FKDEVDIYALSQEVCE--CPGRKRYEAFGGITLTNSEK--KEIKVPVNWWDKSKQ--P 146

Qy 140 IDKVKTSKKEVTVQELDLQARHYLHGKFGLYNSDSFGKQVQGLIVFHSSEGSTVSYDL- 198
Db 147 PMFITVKNPKVTAQEVDDIKVRKLLIKKYDIYNNRE--QKYSKGTVTLDLNSGKDIVEFDLY 204

Qy 199 FDAQGYPTDLLRIYRDNTTISSTLSLSL 228
Db 205 YFGNGDF-NSMLKIYSNNERIDSTQFHVYD 233

RESULT 11
SPEC_STRPY
ID SPG_STRPY STANDARD; PRT; 234 AA.
AC Q9X5C7;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exotoxin type G precursor (SPE G).
GN SPG OR SPY0212.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RC STRAIN=M1;
RX MEDLINE=99093428; PubMed=9874566;
RA Proft T., Moffatt S.L., Berkahn C.J., Fraser J.D.;
RT "Identification and characterization of novel superantigens from
RL Streptococcus pyogenes."
RL J. Exp. Med. 189:89-102(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
CC -I- SUBUNIT: Binds to major histocompatibility complex class II beta
CC chain.
CC -I- DISEASE: Mitogenic for human peripheral blood lymphocytes.
CC -I- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF124499; AAD30988.1; -.
CC EMBL: AE006489; AAK33303.1; -.
CC HSSP: P13380; 1AN8.
CC InterPro: IPR001961; Staph/Strep_toxin.
CC Pfam: PF01123; Staph_Strp_toxin; 1.
CC Pfam: PF02876; Staph_Strp_tox_C; 1.
CC PRINTS: PR00279; BACTRLTOXIN.
CC PROSITE: PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.
CC PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
CC Toxin; Signal; Complete proteome.
CC SIGNAL 1 24 POTENTIAL.
CC CHAIN 25 234 EXOTOXIN TYPE G.
CC SEQUENCE 234 AA; 27262 MW; 49525C49E4BA2052 CRC64;

Query Match 15.18; Score 183.5; DB 1; Length 234;
Best Local Similarity 27.68; Pred. No. 4.4e-09;
Matches 56; Conservative 37; Mismatches 81; Indels 29; Gaps 7;

Qy 33 NSKAITSEKSDAQFLNTLLFKGFTTGHWPYNLLVDLGSTAAATSEYEGSSVDLYGAY 92
Db 57 NSIHNTKQRSECI-----YVDSIVSLGIT--DQFIKGDKVDVFGLPY 99

Qy 93 GYCAGGTPNKACMGVTLHDNNRLTEKKVPTNLWIDGKQTTPVTDKVKTSKKEVTV 152
Db 100 NF-----SPPYVDNIYGGIVKHSNQGKSLQFVGTLNQDGKTYLPSEAVRIKKQFTL 153

Qy 153 QELDLQARHYLHGKFGLYNSDSFGKQVQGLIVFHSSEGSTVSYDLFDAQGY--PDFLL 210
Db 154 QEFDFKIRKFLMEKINYDSES---RYTSGSLFLATKDSKHVEVDLFNKDKLLSRDSFF 210

Qy 211 RIYRDNTTISSTLS-LSLYLT 232
Db 211 KRYKDNKIFNSEEISHFDIYLT 233

RESULT 12
SPEC_STRPY
ID SPG_STRPY STANDARD; PRT; 235 AA.
AC P13380;
DT 01-JAN-1990 (Rel. 13, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exotoxin type C precursor (SPE C).
```

GN SPEC OR SPY0711.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 28-52.
RC STRAIN-T18P / MGAS 1585;
RX MEDLINE=88314303; PubMed=3045005;
RA Goshorn S.C., Schliefert P.M.;
RT "Nucleotide sequence of streptococcal pyrogenic exotoxin type C";
RL Infect. Immun. 56:2518-2520(1988).
RN [2]
RP REVISIONS TO 21-26.
RC STRAIN-T18P / MGAS 1585;
RX MEDLINE=92363541; PubMed=1500157;
RA Kapur V., Nelson K., Schliefert P.M., Selander R.K., Musser J.M.;
RT "Molecular population genetic evidence of horizontal spread of two
RT alleles of the pyrogenic exotoxin C gene (speC) among pathogenic
RT clones of Streptococcus pyogenes";
RL Infect. Immun. 60:3513-3517(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferrer J.J., McShan W.M., Ajdic D.J., Savic D.J., Lyon K.,
RA Prlmeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus
RT pyogenes";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 30-235.
RX MEDLINE=97397352; PubMed=9253413;
RA Rousset A., Anderson B.F., Baker H.M., Fraser J.D., Baker E.N.;
RT "Crystal structure of the streptococcal superantigen SPE-C:
RT dimerization and zinc binding suggest a novel mode of interaction
RT with MHC class II molecules";
RL Nat. Struct. Biol. 4:635-643(1997).
CC -!- SUBUNIT: Binds to major histocompatibility complex class II beta
CC chain
CC -!- DISEASE: THE STREPTOCOCCAL PYROGENIC TOXINS A, B, AND C ARE
CC THE CAUSATIVE AGENTS OF THE SYMPTOMS ASSOCIATED WITH SCARLET
CC FEVER, HAVE BEEN ASSOCIATED WITH STREPTOCOCCAL TOXIC SHOCK-LIKE
CC DISEASE AND MAY PLAY A ROLE IN THE EARLY EVENTS OF RHEUMATIC
CC FEVER.
CC -!- MISCELLANEOUS: THIS TOXIN SEEMS TO BE CODED BY A BACTERIOPHAGE.
CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M35514; AA27017.1; ALT_SEQ.
CC EMBL: M97156; AAB59091.1; -.
CC EMBL: M97157; AAB59092.1; -.
CC EMBL: AE006523; AAK33664.1; -.
CC PIR: A30509; A30509.
CC PDB: 1AN8; 29-APR-98.
CC InterPro: IPR001961; Staph/Strep_toxin.
CC Pfam: PF01123; Staph_Strep_toxin; 1.
CC Pfam: PF02876; Staph_Strep_tox_C; 1.
CC PRINTS: P00279; BACTRLTOXIN.
CC PROSITE: PS00277; STAPH_STREP_TOXIN_1; 1.
CC PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
CC Toxin; Signal; 3D-structure: Complete proteome.
CC SIGNAL 1 27

FT CHAIN 28 235 EXOTOXIN TYPE C.
FT CONFLICT 53 N -> D (IN REF. 1).
SQ SEQUENCE 235 AA; 27371 MW; 070534ABB952C1E0 CRC64;
Query Match 12.2%; Score 148.5; DB 1; Length 235;
Best Local Similarity 25.6%; Pred. No. 5.3e-06;
Matches 58; Conservative 39; Mismatches 85; Indels 45; Gaps 11;
QY 34 SKAITSSEKSADQFLTNTLLFKGFFTHPEW-YNDLLVDLGLSTAA-----TSEYEGSS---- 84
DB 22 SPLIKSDSKDISNVKSDLLYA--YTIPTDYKNCRVNFTHTLLNIDITQYRGKDYIIS 79
QY 85 -----VDLYGAYGYQCAGGTPNKTACMGVTVLHDNRLTEKKVPIN 128
DB 80 SEMSYEASQKFRKRDHVDVFLGLIILNSHTG-----EYIYGGITPAQNNKVNH--KLLGN 132
QY 129 LMTDGRQTTVPIDKVKTSKKEVTQELDLQARHLYHGKGLYNSDS--FGGKVQRGGLIVF 186
DB 133 LFISGESQNLNNKIILEKDIYTFQEIFKIKYLMNDKIYDATSPYVSGRIETG---- 188
QY 187 HSSEGSTVSVDLFDFA--QGQYPTLLRIYRDNNTTISSTLS-ISLYL 230
DB 189 -TKDGKHEQIDLFDSPEGTRSDIFAK-YKDNRIINKNKFNFSHDIYL 233
RESULT 13
TSST_STAAU
ID TSST_STAAU STANDARD; PRT; 234 AA.
AC P06886;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Toxic shock syndrome toxin-1 precursor (TSST-1).
GN TST.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=87057222; PubMed=3782090;
RA Blonster-Hautamaa D.A., Kreiswirth B.N., Kornblum J.S., Novick R.P.,
RA Schliefert P.M.;
RT "The nucleotide and partial amino acid sequence of toxic shock
RT syndrome toxin-1";
RL J. Biol. Chem. 261:15783-15786(1986).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=94150598; PubMed=8107781;
RA Acharya K.R., Passalacqua E.F., Jones E.Y., Harlos K., Stuart D.I.,
RA Brehm R.D., Tranter H.S.;
RT "Structural basis of superantigen action inferred from crystal
RT structure of toxic-shock syndrome toxin-1";
RN Nature 367:94-97(1994).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RX MEDLINE=94092653; PubMed=8268150;
RA Prasad G.S., Earhart C.A., Murray D.L., Novick R.P., Schliefert P.M.,
RA Ohlendorf D.H.;
RT "Structure of toxic shock syndrome toxin 1";
RL Biochemistry 32:13761-13766(1993).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.07 ANGSTROMS).
RX MEDLINE=96319751; PubMed=8759320;
RA Papageorgiou A.C., Brehm R.D., Leonidas D.D., Tranter H.S.,
RA Acharya K.R.;
RT "The refined crystal structure of toxic shock syndrome toxin-1 at
RT 2.07-A resolution";
RL J. Mol. Biol. 260:553-569(1996).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.05 ANGSTROMS).
RX MEDLINE=97337442; PubMed=9194182;
RA Prasad G.S., Radhakrishnan R., Mitchell D.T., Earhart C.A.,
RA Dinges M.M., Cook W.J., Schliefert P.M., Ohlendorf D.H.;

RT "Refined structures of three crystal forms of toxic shock syndrome
RL toxin-1 and of a tetramutant with reduced activity.";
RN [6]

RP X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS) OF MUTANTS.

RX MEDLINE=98254504; PubMed=9585531;

RA Earhart C.A., Mitchell D.T., Murray D.L., Pinheiro D.M., Matsumura M.,

RA Schlievert P.M., Ohlendorf D.H.;

RT "Structures of five mutants of toxic shock syndrome toxin-1 with

RT reduced biological activity.";

RL Biochemistry 37:7194-7202(1998).

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- DISEASE: THIS TOXIN IS RESPONSIBLE FOR THE SYMPTOMS OF TOXIC

CC SHOCK SYNDROME.

CC -!- SIMILARITY: BELONGS TO THE STAPHYLOCOCCAL/STREPTOCOCCAL TOXIN

CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: J02615; AAA26682.1; .

DR PIR: A24606; XCSAS1

DR PDB: 2TSS; 24-DEC-97.

DR PDB: 3TSS; 24-DEC-97.

DR PDB: 4TSS; 24-DEC-97.

DR PDB: 5TSS; 24-DEC-97.

DR PDB: 10IL; 12-AUG-97.

DR PDB: 20IL; 12-AUG-97.

DR PDB: 1AW7; 18-NOV-98.

DR PDB: 1TS2; 16-DEC-98.

DR PDB: 1TS3; 16-DEC-98.

DR PDB: 1TS4; 16-DEC-98.

DR PDB: 1TS5; 16-DEC-98.

DR InterPro: IPR001961; Staph/Strep_toxin.

DR Pfam: PF01123; Staph_Strep_toxin; 1.

DR Pfam: PF02876; Staph_Strep_tox_C; 1.

DR PRINTS: PR01501; TOXICSTOXIN.

DR PROSITE: PS00277; STAPH_STREP_TOXIN_1; FALSE_NEG.

DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.

KW Toxin; Superantigen; Signal; 3D-structure.

FT SIGNAL

FT CHAIN 1 40

SQ SEQUENCE 234 AA; 26306 MW; E95789FF9A1D7AB4 CRC64;

Query Match 8.2%; Score 100; DB 1; Length 234;

Best Local Similarity 24.5%; Pred. No. 0.099; Mismatches 33; Indels 54; Gaps 12;

Matches 48; Conservative 33; Mismatches 61; Indels 54; Gaps 12;

QY 25 NLKQI---YYNSKAITSE-----KSAQFLNTLLFKGFTTGHWPYNDLL-- 68

Db 45 NIKDLLWYSSGSDFTNSEVLDNLSGSRKNTDGSII-SLIIFPS-----PYSPATK 98

QY 69 ---VDLGSTAATSEYGGSDVLYGAYGQCAGGTPNKTACMYGGVTLHDNNRLTEKKV 125

Db 99 GEKVLD-NTKRRTKKSQHTS---EGTYIHFIQISGVT-----NTEKLPPTIEL 140

QY 126 PINLWIDGKQTVPIDKVTKSKKVTQVQLDQARHLYLHGKFGLY-NSDSFGG--KVQRG 182

Db 141 PLKRVKHGKDSPLKYGKPKFDKQKLAISTLDPEIRHQLTQHGLYRSSDKTGGYKWT-- 197

QY 183 LIVFHSSEGSTVSVDL 198

Db 198 -----MNDGSTVQSD 208

RESULT 14

R19E_THEVO

ID R19E_THEVO STANDARD; PRT; 150 AA.
AC Q97CU4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S19E.
GN RPS19E OR TV0007 OR TVG0008143.
OS Thermoplasma volcanium.
OC Archaea; Euryarchaeota; Thermoplasmata; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=50339;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=GSSI / DSM 4299 / JCM 9571;
RX MEDLINE=20570466; PubMed=11121031;
RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
RA Kawashima T., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,
RA Nunoshima T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;
RT "Archaeal adaptation to higher temperatures revealed by genomic
RT sequence of Thermoplasma volcanium".
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
CC -!- SIMILARITY: BELONGS TO THE S19E FAMILY OF RIBOSOMAL PROTEINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AP000991; BAB59149.1; .
DR InterPro: IPR001266; Ribosomal_S19E.
DR Pfam: PF01090; Ribosomal_S19e; 1.
DR PROSITE: PS00628; RIBOSOMAL_S19E; FALSE_NEG.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 150 AA; 17063 MW; FD5881CF684EB415 CRC64;

Query Match 7.8%; Score 95.5; DB 1; Length 150;
Best Local Similarity 29.5%; Pred. No. 0.14; Mismatches 17; Indels 1; Gaps 1;
Matches 23; Conservative 17; Mismatches 37; Indels 1; Gaps 1;

QY 116 NNRLTEERKKVPINLWIDGKQTVPIDKVTKSKKVTQVQLDQARHLYLHGKFGLYN-SDS 174
Db 17 SGKLEKKIKPEPNVVKVTKGVSKEKPPLODDWLYVRAASMLRKLKLYINGLYGISMSE 76

QY 175 FGGKVORGLVIFHSSEGS 192
Db 77 YGGKVDGRGSKRYHAASGS 94

RESULT 15
P115_MYCHR
ID P115_MYCHR STANDARD; PRT; 979 AA.
AC P41508;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE P115 protein.
OS Mycoplasma hyorhinis.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2100;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91138990; PubMed=1825306;
RA Notariccola S.M., McIntosh M.A., Wise K.S.;
RT "A Mycoplasma hyorhinis protein with sequence similarities to
RT nucleotide-binding enzymes".
RL Gene 97:77-85(1991).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- DOMAIN: CONSISTS OF TWO PUTATIVE CENTRAL COILED-COIL REGIONS
CC FLANKED BY PUTATIVE GLOBULAR REGIONS AT THE N- AND C-TERMINUS.
CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. SIMILAR TO OTHER MYCOPLASMA

```
CC p115.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M34956; AAA25423.1; -
CC PIR; JQ0894; JQ0894; -
CC InterPro; IPR003439; ABC_transportr.
CC InterPro; IPR005289; GTP-binding_dom.
CC InterPro; IPR003405; SMC_C.
CC InterPro; IPR003395; SMC_N.
CC Pfam; PF02463; SMC_N; 1.
CC Pfam; PF02483; SMC_C; 1.
CC TIGRfams; TIGR00650; MG442; 2.
KW ATP-binding; Coiled coil.
FT NP_BIND 32 39 ATP (POTENTIAL).
FT DOMAIN 169 224 COILED COIL (POTENTIAL).
FT DOMAIN 231 400 COILED COIL (POTENTIAL).
FT DOMAIN 569 821 COILED COIL (POTENTIAL).
FT DOMAIN 884 912 ALA/ASP-RICH (DA-BOX).
SQ SEQUENCE 979 AA; 110566 MW; 30D51C5B56280F4 CRC64;

Query Match 7.6%; Score 93; DB 1; Length 979;
Best Local Similarity 23.3%; Pred. No. 2.4;
Matches 55; Conservative 41; Mismatches 92; Indels 48; Gaps 9;

QY 17 ELGGTALGNLKOIYYNSKATSSKSDQ-----FLTNTLLFKGFFTG 61
DB 56 QLRLG---NMDDVIFAGSKTVKPOEKAMVKLTPKNDIAETETKQIFTSRLLRKGGTNE 112
QY 62 PWYNDLLV---DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACHY---GGVTLHD 115
DB 113 YFYNDOPVRYKDKLNLAVERGSKSLAISQGTISEIAEATPEQRKAVIEEAAGTSKYK 172
QY 116 NNRLTEKKVPINLWDIGKQTTPIDKVTSSKKEV---TVQELDLQ---ARHYLHGKFGLY 170
DB 173 LDKEEAQKLL-----IFTNDAIDKLOGAIKELERQVNSLDKQASKAKIYLEKSKALE 224
QY 171 NSDSFGKGVQRLV-----FHSSEGTSVSYDLFDAGQVPTLLRIYRONTTISST 222
DB 225 S-----VEVGLIVNLAFFNEKLNNTSLLEVEQQRNDLELNIQTVESSISQT 273

RESULT 16
DPOL_METJA STANDARD; PRT; 1634 AA.
AC O58295;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE DNA polymerase (EC 2.7.7.7) [Contains: Mja pol-1 intein; Mja pol-2 intein].
GN POL OR MJ0885.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcales; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
```

```
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL jannaschii.";
RL Science 273:1058-1073(1996).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate -> N diphosphate
CC + (DNA)(N).
CC -!- PTM: THIS PROTEIN UNDERGOES A PROTEIN SELF SPLICING THAT INVOLVES
CC A POST-TRANSLATIONAL EXCISION OF THE INTERVENING REGION (INTEIN)
CC FOLLOWED BY PEPTIDE LIGATION (POTENTIAL).
CC -!- SIMILARITY: BELONGS TO THE DNA POLYMERASE TYPE-B FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U67532; AAB98889.1; -
CC HSSP; P56689; LTGO.
CC TIGR; MJ0885; -
CC InterPro; IPR002064; DNA_pol_B.
CC InterPro; IPR003586; Hedgehog_hintc.
CC InterPro; IPR003587; Hedgehog_hintc.
CC InterPro; IPR002203; Intein.
CC InterPro; IPR004042; Intein_endonuc.
CC InterPro; IPR004578; Pol2.
CC Pfam; PF00136; DNA_pol_B; 3.
CC Pfam; PF03104; DNA_pol_B_exo; 1.
CC PRINTS; SM00106; DNAPOLB.
CC SMART; SM00305; Hintc; 2.
CC SMART; SM00306; Hintc; 2.
CC SMART; SM00486; POLBC; 1.
CC TIGRfams; TIGR00592; pol2; 1.
CC PROSITE; PS00116; DNA_POLYMERASE_B; 1.
CC PROSITE; PS50818; INTEIN_C-TER; 2.
CC PROSITE; PS50819; INTEIN_ENDONUCLEASE; 2.
CC PROSITE; PS50817; INTEIN_N-TER; 2.
KW Transferase: DNA-directed DNA polymerase; DNA replication;
KW DNA-binding; Autocatalytic cleavage; Protein splicing;
KW Complete proteome.
FT CHAIN 1 425 POL, 1ST PART (POTENTIAL).
FT CHAIN 426 794 MJA POL-1 INTEIN (POTENTIAL).
FT CHAIN 795 882 POL, 2ND PART (POTENTIAL).
FT CHAIN 883 1358 MJA POL-2 INTEIN (POTENTIAL).
FT CHAIN 1359 1634 POL, 3RD PART (POTENTIAL).
SQ SEQUENCE 1634 AA; 191708 MW; 84A1FAFAB1F97DDD CRC64;

Query Match 7.5%; Score 91; DB 1; Length 1634;
Best Local Similarity 21.6%; Pred. No. 6.8;
Matches 50; Conservative 36; Mismatches 79; Indels 66; Gaps 10;

QY 12 LRKSELOGTALGNLKOIYYNSKATSSKSDQFLNTLLFKGFFTGHPWNLLVDL 71
DB 614 VRRKKTKAITLGCARKDIYKTEIL---KNKRYLPNAIL-RGFEFGDYVNTVRRV 669
QY 72 GSTAATSEYE----GSSVDLYGAYGYQCAGGTPNKTACMYGVTLHNNRLTEKKVP 126
DB 670 VYNOGTNNYDKIKFIASLLDLRLGIKYSF-----YTSYEERKKLRYV 713
QY 127 INLWDIGKQTTPIDRKVTS-----RKEVTVQELDLQARHYLHGKFGLYNSD----- 173
DB 714 IEIFSKG-----DLIKFSILISFISRRKNLLNEIRKTKLYKIGDYGYDLDDCVS 766
QY 174 --SFGKVV-----QRGLIVFHSSEGS-TVSYDLFDAGQVPTPL 209
DB 767 LESYKGEVYDLTLEGRPPYFANGILTHNSLYPSIIISYNI-----SPDTL 811

RESULT 17
ID R19E_THEAC STANDARD; PRT; 150 AA.
```



```

RP SEQUENCE OF 318-393 FROM N.A.
RX MEDLINE=94008528; PubMed=8404528;
RA Whittaker C.A., Desimone D.W.;
RT "Integrin alpha subunit mRNAs are differentially expressed in early
RT Xenopus embryos";
RL Development 117:1239-1249(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRONECTIN.
CC IT RECOGNIZES THE SEQUENCE R-G-D IN ITS LIGANDS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
CC DISULFIDE BOND. ALPHA-5 ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL; U12683; AAA99668.1; -.
CC EMBL; LI0191; AAA16249.1; -.
CC HSSP; P06756; 1JY2.
CC InterPro; IPR000413; Integrin_alpha.
CC Pfam; PF00357; Integrin_A; 1.
CC Pfam; PF01839; FG-GAP; 5.
CC PRINTS; PRO1185; INTEGRINA.
CC SMART; SM00191; Int_alpha; 5.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat.
FT SIGNAL 1 32 POTENTIAL.
FT CHAIN 33 1050 INTEGRIN ALPHA-5.
FT CHAIN 33 932 INTEGRIN ALPHA-5 HEAVY CHAIN (POTENTIAL).
FT CHAIN 933 1050 INTEGRIN ALPHA-5 LIGHT CHAIN (POTENTIAL).
FT DOMAIN 33 996 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 997 1022 POTENTIAL.
FT DOMAIN 1023 1050 CYTOPLASMIC (POTENTIAL).
FT REPEAT 48 110 FG-GAP 1.
FT REPEAT 120 188 FG-GAP 2.
FT REPEAT 189 246 FG-GAP 3.
FT REPEAT 259 312 FG-GAP 4.
FT REPEAT 313 373 FG-GAP 5.
FT REPEAT 379 438 FG-GAP 6.
FT REPEAT 442 494 FG-GAP 7.
FT CA_BIND 324 332 POTENTIAL.
FT CA_BIND 390 398 POTENTIAL.
FT CA_BIND 454 462 POTENTIAL.
FT SITE 1025 1029 GFGR MOTIF.
FT DISULFID 90 99 BY SIMILARITY.
FT DISULFID 145 166 BY SIMILARITY.
FT DISULFID 182 195 BY SIMILARITY.
FT DISULFID 502 513 BY SIMILARITY.
FT DISULFID 519 575 BY SIMILARITY.
FT DISULFID 636 642 BY SIMILARITY.
FT DISULFID 708 721 BY SIMILARITY.
FT DISULFID 862 910 INTERCHAIN (BY SIMILARITY).
FT DISULFID 917 922 BY SIMILARITY.
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 507 507 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 515 515 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 521 521 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 600 600 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 649 649 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 714 714 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 763 763 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 861 861 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1050 AA; 115961 MW; 10ED961535B8D918 CRC64;

Query Match 6.9%; Score 84; DB 1; Length 1050;
Best Local Similarity 21.9%; Pred. No. 16;
Matches 60; Conservative 38; Mismatches 94; Indels 82; Gaps 14;

QY 12 LRKSELO-GTALGNLKOIYYNSKAITS-SEKSAOFLTNTLLFKGFFTHGHPWYNDLLV 69
DB 12 LRKSELO-GTALGNLKOIYYNSKAITS-SEKSAOFLTNTLLFKGFFTHGHPWYNDLLV 69
DB 242 LEYKKOMQTRQAASYSYDYSFYGYSVAVGEFSEDAEDFVV-----GYPKNGTYGYVTI 295
QY 70 DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTA---CMYGGVTLHD---NNRITTEK 123
DB 296 -LNGTDLRLSYNFSGEGM-ASYFGYSVSATDLNSDGLDLDLLICAPLFMDRTHGRVQEVG 353
QY 124 KVPINLWIDGKQTPVDKVKTSKKEVTVOELDLQARHYLHGKFG-----LYNSDS---- 174
DB 124 KVPINLWIDGKQTPVDKVKTSKKEVTVOELDLQARHYLHGKFG-----LYNSDS---- 174
DB 354 RVVYVLQGDHMESTPHL--ILTGME-----YGRFGSSIASGLDLDQDGFN 397
QY 175 -----FGKQVORGLI-VFHSSEGSTVSVDLFDAGQ----- 204
DB 398 DIAIGAPFGGEAQRGAVFINGQPGGVDSKPSQVLOGMGSSQSPFFGLSTRGHDLG 457
QY 205 --YPTDLLRIYRDNNTI-----SSTSLSTS 227
DB 458 NGYPDLIVGAFGVDTTLVYRGRPIIHASASLSIS 491

RESULT 23
PRTA:ASPNG
ID PRTA:ASPNG STANDARD; PRT; 282 AA.
AC P24665;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Aspergillopepsin II precursor (EC 3.4.23.19) (Acid protease A)
DE (Protease A).
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Var. Macrosporus;
RX MEDLINE=92011747; PubMed=1918060;
RA Inoue H., Kimura T., Makabe O., Takahashi K.;
RT "The gene and deduced protein sequences of the zymogen of Aspergillus
RT niger acid proteinase A.";
RL J. Biol. Chem. 266:19484-19489(1991).
RN [2]
RP SEQUENCE OF 60-98 AND 110-282.
RC STRAIN=Var. Macrosporus;
RX MEDLINE=92011746; PubMed=1918059;
RA Takahashi K., Inoue H., Sakai K., Kohama T., Kitahara S.,
RA Takashima K., Tanji M., Athauda S.B.P., Takahashi T., Akanuma H.,
RA Mamiya G., Yamasaki M.;
RT "The primary structure of Aspergillus niger acid proteinase A.";
RL J. Biol. Chem. 266:19480-19483(1991).
CC -!- CATALYTIC ACTIVITY: Preferential cleavage in A chain of insulin:
CC 3-Asn-1-Gln-4, 13-Gly-1-Ala-14, and 26-Tyr-1-Thr-27.
CC -!- SUBUNIT: HETERODIMER OF TWO NONCOVALENTLY BOUND LIGHT AND HEAVY
CC CHAINS.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A4.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)

```

```
CC EMBL; M68871; AAA32687.1; -.
DR PIR; A41025; A41025.
DR MEROPS; A04.002; -.
DR InterPro; IPR000250; AspartaseA4.
DR Pfam; PF01828; Peptidase_A4; 1.
DR PRINTS; PR00977; SCYTLDPSTASE.
DR ProDom; PD018627; SCYTLDPSTASE_A4; 1.
KW Hydrolase; Aspartyl protease; Zymogen; Signal.
FT SIGNAL 1 18
FT PROPEP 19 59
FT CHAIN 60 98
FT PROPEP 99 109
FT CHAIN 110 282
FT MOD_RES 110 110
FT DISULFID 115 139
FT DISULFID 127 210
SQ SEQUENCE 282 AA; 29887 MW; 4EA727F9AE33F72A CRC64;

Query Match 6.9%; Score 83.5; DB 1; Length 282;
Best Local Similarity 19.2%; Pred. No. 3.5;
Matches 52; Conservative 34; Mismatches 102; Indels 83; Gaps 10;

QY 13 RKSELOQTALGNLKOIYYNSKAITSEKSDAQFUTNLLFKGFTTGHWPYNLDLLVLDG 72
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 28 RKEARAAGKRHSNPPYIPGSDKEILKLNCTTNEEYSSN-----WAGAVLIGDG 75

QY 73 STAANTSEYEGSSVDL-----YGAYGY-----QCAG-----GTPNKTACMYGG 110
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 76 YTKVTGFTVPSVAGSSGSGGYGGYGNKRNKROSEECASAWVGIDGDTCTAILQRT 135

QY 111 VTLHDNRNLTETKVPINLWI-----DGKQTTVPIDKVKTSKKEVTVOE 154
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 136 VDF-----CYEGQTSYDANYEWPYDYDFSDITISEGDSIKVTVEATSKSGSATVEN 190

QY 155 LDL--QARHYLHGKF-----GLYNSDFGKGVORGLIVFHSS-----GSTV 194
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 191 LTTGQSVTHTFSGNVEBGLCETNAEWIVDFESGSLVAFADFGSVTFNNAETSGSGTV 250

QY 195 SYDLFDAQOQYPTLLRIYRDNTTISSTLS 225
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 251 -----GPSDATVMDIEQDGSVLTFETSVS 273

RESULT 24
YB70_HAEIN
ID YB70_HAEIN STANDARD; PRT; 328 AA.
AC Q57527; O05047;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical paba-like protein H1170.
GN H1170.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae Rd."
RL Science 269:496-512(1995).
```

```
CC -!- SIMILARITY: TO THE C-TERMINAL OF PARA-AMINOBENZOATE SYNTHASE
CC COMPONENT I.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U32797; AAC22834.1; -.
DR TIGR; H11170; -.
DR InterPro; IPR000350; Chorismate_bind.
DR Pfam; PF00425; chorismate_bind; 1.
DR ProDom; PD000779; Chorismate_bind; 1.
DR Hypothetical protein; Complete proteome.
KW SEQUENCE 328 AA; 37734 MW; IC02BCD0088957B0 CRC64;

Query Match 6.8%; Score 82.5; DB 1; Length 328;
Best Local Similarity 27.3%; Pred. No. 5.2;
Matches 24; Conservative 16; Mismatches 35; Indels 13; Gaps 3;

QY 136 TTVPIDKVKTSKKEVT---VQELDLQARHYLHGKFLYNSDFSFGKVGQRLI-----V 185
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 237 TLLPAGSISGAPKKTQIIQOAEKQKRGYTGIFGIDGKTLQSAVAIRFISQVDEKEY 296

QY 186 FHSSEGSSTVSVDLFDAAQOQYPTLLRIY 213
||:| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 297 FHSGGGITIH---SNAQDEYEELLEKVV 321

RESULT 25
YB70_BACSU
ID YB70_BACSU STANDARD; PRT; 455 AA.
AC P25152;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical peptidase ywad precursor (EC 3.4.11.-).
GN YWAD OR IPA-8R.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=95020537; PubMed=7934828;
RA Glaser P., Kunst F., Arnaud M., Coudart M.P., Gonzales W.,
RA Hullo M.F., Ionescu M., Lubochinsky B., Marcelino L., Moszer I.,
RA Presecan E., Santana M., Schneider E., Schweizer J., Vertes A.,
RA Rapoport G., Danchin A.;
RT "Bacillus subtilis genome project: cloning and sequencing of the 97
RT kb region from 325 degrees to 333 degrees."
RL Mol. Microbiol. 10:371-384(1993).
RN [2]
RP SEQUENCE OF 1-68 FROM N.A.
RC STRAIN=168;
RX MEDLINE=92216127; PubMed=1806041;
RA Glaser P., Kunst F., Debarbouille M., Vertes A., Danchin A.,
RA Dedonder R.;
RT "A gene encoding a tyrosine tRNA synthetase is located near sacs in
RT Bacillus subtilis."
RL DNA Seq. 1:251-261(1991).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC EMBL: X73124; CAA51564.1; -
DR EMBL: X52480; CAA36725.1; -
DR EMBL: Z99123; CAB15873.1; -
DR PIR: S16427; S16427
DR PIR: S39663; S39663
DR MEROPS: M28; UPB; -
DR Subtilisin; BG10554; ywad.
DR InterPro: IPR003137; PA.
DR Pfam: PF02225; PA; 1.
KW Hypothetical protein; Hydrolase; Aminopeptidase; Metalloprotease;
KW Zinc; Signal; Complete proteome.
FT SIGNAL 1 31
FT CHAIN 32 455 HYPOTHETICAL PEPTIDASE YWAD.
SQ SEQUENCE 455 AA; 49450 MW; 89BE6A6EEB0CCE18 CRC64;

Query Match 6.8%; Score 82.5; DB 1; Length 455;
Best Local Similarity 22.0%; Pred. No. 7.9; Mismatches 67; Indels 53; Gaps 6;
Matches 41; Conservative 25;

QY 1 SEKSEINERKDLRKSELOCTALGNLKOIYYNSKAITSSEKSADQF-----LTNTLLFK 55
   | | | | | : | | | | | : | | | | | : | | | | |
Db 53 SOLSEALGPR-----IAGTAEKKSALLIASSMRKLLDKVKVQRENIPDRLEGLTSSA 105
   | | | | | : | | | | | : | | | | | : | | | | |
QY 56 GFTTGHFWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQCAGGTGNKTCMYGGVTLHD 115
   | | | | | : | | | | | : | | | | | : | | | | |
Db 106 G-----RDILLQAASGSAPEEQGLTAPLYNAGLGQY-----KDFADAKGKIALIS 152
   | | | | | : | | | | | : | | | | | : | | | | |
QY 116 NNRLT-----EKKVPINLMDGKQITVPTDKVK-----TSK 147
   | | | | | : | | | | | : | | | | | : | | | | |
Db 153 RGLTTYEKAKNAEAGAKAVIYNKRESLVPTNLSGNKGVIPVYGIKKEDGEALTOQ 212
   | | | | | : | | | | | : | | | | | : | | | | |
QY 148 KEVTQ 153
   | | | | | : | | | | | : | | | | | : | | | | |
Db 213 KEATLK 218

RESULT 26
NUCD_ECOLI STANDARD; PRT; 600 AA.
AC P3359; P33600; P78089; P78309;
DT 01-FEB-1994 (Rel. 28, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE NADH dehydrogenase I chain C/D (EC 1.6.5.3) (NADH-ubiquinone
DE oxidoreductase chain 3/4) (NUO3/NUO4).
GN NUOC OR NUOCD OR NUOD OR B2286.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / AN387;
RX MEDLINE=93389724; PubMed=7690854;
RA Weidner U., Geier S., Ptock A., Friedrich T., Leif H., Weiss H.;
RT "The gene locus of the proton-translocating NADH: ubiquinone
RT oxidoreductase in Escherichia coli. Organization of the 14 genes and
RT relationship between the derived proteins and subunits of
RT mitochondrial complex I";
RL J. Mol. Biol. 233:109-122(1993).
RN [2]
RP REVISIONS.
RC STRAIN=K12 / AN387;
RA Weidner U.;
RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

```

```

RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RT Science 277:1453-1474(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshina T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,
RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT -K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN [5]
RP SEQUENCE OF 179-600 FROM N.A.
RX MEDLINE=94209210; PubMed=8157582;
RA Pruss B.M., Nels J.M., Park C., Wolfe A.J.;
RT "Mutations in NADH:ubiquinone oxidoreductase of Escherichia coli
RT affect growth on mixed amino acids.";
RL J. Bacteriol. 176:2143-2150(1994).
CC -!- FUNCTION: TRANSFER OF ELECTRONS FROM NADH TO THE RESPIRATORY
CC CHAIN. THE IMMEDIATE ELECTRON ACCEPTOR FOR THE ENZYME IS BELIEVED
CC TO BE UBIQUINONE. DOES COUPLE THE REDOX REACTION TO PROTON
CC TRANSLOCATION AND THUS CONSERVES THE REDOX ENERGY IN A PROTON
CC GRADIENT.
CC -!- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -!- SUBUNIT: COMPOSED OF 13 DIFFERENT SUBUNITS. SUBUNITS NUOCD, E,
CC F, AND, G CONSTITUTE THE PERIPHERAL SECTOR OF THE COMPLEX.
CC -!- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE COMPLEX I 30
CC KDA SUBUNIT FAMILY.
CC -!- SIMILARITY: IN THE C-TERMINAL SECTION; BELONGS TO THE COMPLEX I 49
CC KDA SUBUNIT FAMILY.
CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 175.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68301; CAA48362.1; ALT_FRAME.
DR EMBL: X68301; CAA48363.1; ALT_FRAME.
DR EMBL: AE000317; AAC75346.1; -.
DR EMBL: D90859; BAA16115.1; -.
DR EMBL: D90860; BAA16120.1; ALT_INIT.
DR EMBL: L25055; AAA03535.1; ALT_INIT.
DR PIR: S37060; S37060.
DR PIR: S38312; S38312.
DR EcoGene; EG12084; nuoc.
DR InterPro: IPR001268; Complex1_30K.
DR InterPro: IPR001135; Oxidored_49kDa.
DR Pfam; PF00329; complex1_30Kd; 1.
DR Pfam; PF00346; complex1_49Kd; 1.
DR ProDom; PD001581; Complex1_30K; 1.
DR ProDom; PD001581; Complex1_30K; 1.
DR PROSITE; PS00542; COMPLEX1_30K; 1.
DR PROSITE; PS00535; COMPLEX1_49K; 1.
KW Oxidoreductase; NAD; Ubiquinone; Multifunctional enzyme;
KW Complete proteome.
FT DOMAIN 1 190 NADH DEHYDROGENASE I CHAIN C.
FT DOMAIN 214 600 NADH DEHYDROGENASE I CHAIN D.
FT CONFLICT 370 370 H -> D (IN REF. 5).
FT CONFLICT 413 416 AYGA -> PWAR (IN REF. 1 AND 5).
FT CONFLICT 495 495 A -> R (IN REF. 5).
SQ SEQUENCE 600 AA; 68694 MW; 0092008D10243A19 CRC64;

Query Match 6.8%; Score 82.5; DB 1; Length 600;

```

```
Best Local Similarity 27.0%; Pred. No. 11;
Matches 51; Conservative 33; Mismatches 76; Indels 29; Gaps 12;

QY 26 LKQIYYNSKAITSSSEKSDQFLTNLLPKGFTGHPWYN-DLLVDLGGTAATSEYEGSS 84
Db 380 LREFDMPKRLASYEKAA---LONTIL-KGRSQGVAAYGAKEALEWGTGAGLRATGID 435
QY 85 VDLYGA--YGYQO-----CAGGTEN-KTACMYGGVTLHDNNRLTEE--KKVPINLW-I 131
Db 436 FDVRKARPYSGYENDEFEPVGGVSDCYTRVLMKVEELRQSLRILEOCLNNMPEGPFA 495
QY 132 DKQQTTPVDIKVKTSEKVTVOELDQARHYLHGKFG--LYNSDSFGG-KVORGLIVFH- 187
Db 496 DHPLTTPP-----PKERTLQIETLTHLQVSMGVPVMPANESQMTAEATKGINSYL 548
QY 188 SSEGSTVSYS 196
Db 549 TSDGSTMSYS 557

RESULT 27
DNAK_STRPY
ID DNAK_STRPY STANDARD; PRT; 607 AA.
AC P95831;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa protein) (HSP70).
GN DNAK OR SPY1760 OR SPYM18_1831.
OS Streptococcus pyogenes, and
OS Streptococcus pyogenes (serotype M18).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314, 186103;
RN [1]
RP SEQUENCE FROM N.A.
RA Rioux C.R., Martin D., Hamel J., Brodeur B.R.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RT Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-WGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barban K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18 group A Streptococcus strains associated with acute rheumatic fever outbreaks.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673(2002).
RN [4]
RP SEQUENCE OF 1-11: 205-215 AND 426-436, AND MASS SPECTROMETRY.
RC STRAIN-JRS4 / Serotype M6;
RA Hogan D.A., Du P., Stevenson T.I., Whitton M., Kilby G.W., Rogers J.,
RA VanBogelen R.A.;
RT "Two-dimensional gel electrophoresis map of Streptococcus pyogenes proteins.";
RC Submitted (MAY-2000) to the SWISS-PROT data bank.
CC !- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).
CC !- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).
CC !- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
-----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U72721; AAB39223.1; -;
DR EMBL; AE006604; AAK34501.1; -;
DR EMBL; AE010089; AAL98349.1; -;
DR HSSP; P04475; IDG4.
DR InterPro; IPR001023; Hsp70.
DR Pfam; PF00012; Hsp70; 1.
DR PRINTS; PR00301; HEATSHOCK70.
DR PRODOM; PD000089; Hsp70; 1.
DR PROSITE; PS00297; HSP70_1; 1.
DR PROSITE; PS00329; HSP70_2; 1.
DR PROSITE; PS01036; HSP70_3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation;
KW Complete proteome.
FT INIT_MET 0
FT MOD_RES 172 172 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 607 AA; 64788 MW; 223BD4C737926F7F CRC64;
Query Match 6.7%; Score 82; DB 1; Length 607;
Best Local Similarity 27.9%; Pred. No. 12;
Matches 39; Conservative 20; Mismatches 57; Indels 24; Gaps 7;
QY 69 VDLGST-AATSEYEGSSVDLYGAYGYCAGGTPNKATCMYGGVTLHD---NNRLTEKK 124
Db 6 IDLTGTSAAVAVLEGTESKIIANPEGNRT---TPSVSEKNGEIIVGDAAKROAVNPET 62
QY 125 VPINLWIDGCKQTPIDIKVTKSKKVTVOELDQARHYLHGKFGLYNSDSFGKQVORGLI 184
Db 63 V---TSIKSKMGT--SEKVSANGKEYTPQEISAMILQYLKG----YAEDYLGEKVEKAVI 113
QY 185 VFHSSEGSTVSVDLFDAAQOQ 204
Db 114 -----TVPAYFENDAAQRQ 125

RESULT 28
PULA_THEMA
ID PULA_THEMA STANDARD; PRT; 843 AA.
AC Q33840;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pullulanase precursor (EC 3.2.1.41) (Alpha-dextrin endo-1,6-alpha-glucosidase) (Pullulan 6-glucanohydrolase).
GN PULA OR TM1845.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogae (class); Thermotogales;
OX Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MSB8 / DSM 3109;
RX MEDLINE=98115241; PubMed=9453151;
RA Bibel M., Brett C., Gossler U., Kriegshaeuser G., Liebl W.;
RT "Isolation and analysis of genes for amylolytic enzymes of the hyperthermophilic bacterium Thermotoga maritima.";
RL FEMS Microbiol. Lett. 158:9-15(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linner K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
```


FT	DOMAIN	856	1442	LAMININ DOMAIN IV 1 (DOMAIN IV B).	FT	DISULFID	1896	1912	BY SIMILARITY.
FT	DOMAIN	1443	1488	LAMININ EGF-LIKE 12.	FT	DISULFID	1915	1930	BY SIMILARITY.
FT	DOMAIN	1489	1532	LAMININ EGF-LIKE 13.	FT	DISULFID	1917	1939	BY SIMILARITY.
FT	DOMAIN	1533	1581	LAMININ EGF-LIKE 14.	FT	DISULFID	1941	1950	BY SIMILARITY.
FT	DOMAIN	1582	1632	LAMININ EGF-LIKE 15.	FT	DISULFID	1953	1968	BY SIMILARITY.
FT	DOMAIN	1633	1642	LAMININ EGF-LIKE 16 (N-TERMINAL).	FT	DISULFID	1971	1986	BY SIMILARITY.
FT	DOMAIN	1643	1831	LAMININ DOMAIN IV 2 (DOMAIN IV A).	FT	DISULFID	1973	1993	BY SIMILARITY.
FT	DOMAIN	1832	1864	LAMININ EGF-LIKE 16 (C-TERMINAL).	FT	DISULFID	1996	2005	BY SIMILARITY.
FT	DOMAIN	1865	1914	LAMININ EGF-LIKE 17.	FT	DISULFID	2008	2022	BY SIMILARITY.
FT	DOMAIN	1915	1970	LAMININ EGF-LIKE 18.	FT	DISULFID	2072	2083	BY SIMILARITY.
FT	DOMAIN	1971	2024	LAMININ EGF-LIKE 19.	FT	DISULFID	2074	2090	BY SIMILARITY.
FT	DOMAIN	2025	2071	LAMININ EGF-LIKE 20.	FT	DISULFID	2092	2101	BY SIMILARITY.
FT	DOMAIN	2072	2118	LAMININ EGF-LIKE 21.	FT	DISULFID	2104	2116	BY SIMILARITY.
FT	DOMAIN	2119	2168	LAMININ EGF-LIKE 22.	FT	DISULFID	2119	2126	BY SIMILARITY.
FT	DOMAIN	2169	2735	DOMAIN II AND I.	FT	DISULFID	2121	2133	BY SIMILARITY.
FT	DOMAIN	2736	2933	LAMININ G-LIKE 1.	FT	DISULFID	2135	2144	BY SIMILARITY.
FT	DOMAIN	2947	3119	LAMININ G-LIKE 2.	FT	DISULFID	2147	2166	BY SIMILARITY.
FT	DOMAIN	3128	3296	LAMININ G-LIKE 3.	FT	DISULFID	2169	2172	INTERCHAIN (PROBABLE).
FT	DOMAIN	3337	3511	LAMININ G-LIKE 4.	FT	DISULFID	2172	2172	INTERCHAIN (PROBABLE).
FT	DOMAIN	3518	3689	LAMININ G-LIKE 5.	FT	CARBOHYD	100	100	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DOMAIN	2205	2257	COILED COIL (POTENTIAL).	FT	CARBOHYD	148	148	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DOMAIN	2330	2464	COILED COIL (POTENTIAL).	FT	CARBOHYD	248	248	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DOMAIN	2604	2621	COILED COIL (POTENTIAL).	FT	CARBOHYD	383	383	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DOMAIN	2639	2705	CELL ATTACHMENT SITE (POTENTIAL).	FT	CARBOHYD	457	457	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SITE	1723	1725	CELL ATTACHMENT SITE (POTENTIAL).	FT	CARBOHYD	485	485	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	SITE	1839	1841	CELL ATTACHMENT SITE (POTENTIAL).	FT	CARBOHYD	905	905	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	305	314	BY SIMILARITY.	FT	CARBOHYD	926	926	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	307	327	BY SIMILARITY.	FT	CARBOHYD	964	964	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	329	338	BY SIMILARITY.	FT	CARBOHYD	1335	1335	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	341	361	BY SIMILARITY.	FT	CARBOHYD	1534	1534	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	364	373	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	366	398	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	401	410	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	413	431	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	434	445	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	436	452	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	454	463	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	466	476	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	500	512	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	502	521	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	523	532	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	535	544	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	547	559	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	549	566	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	568	577	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	580	590	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	593	605	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	595	611	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	613	622	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	625	635	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	638	650	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	640	656	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	658	667	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	670	680	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	683	695	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	685	702	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	704	713	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	716	726	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1443	1455	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1445	1462	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1464	1473	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1476	1486	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1533	1548	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1535	1548	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1555	1555	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1557	1566	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1569	1579	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1582	1594	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1584	1601	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1603	1612	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1615	1630	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1665	1874	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1867	1881	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	DISULFID	1884	1893	BY SIMILARITY.	FT	CARBOHYD	2021	2021	N-LINKED (GLCNAC. .) (POTENTIAL).

Query Match

Best Local Similarity 6.7%; Score 82; DB 1; Length 3718;

Matches 55; Conservative 22; Mismatches 76; Indels 40; Gaps 12;

Qy	58	FTGHPWYNDLL-----VDLGGTAATSEYEGSDVLYGAYGYQACAGTGNKT-ACMY	108
Db	48	FSLHPYFNLAEGARITASATCGEEAPTRSVSRPTEDLYCKLVGPGVAGDPNQIQOY	107
Qy	109	GGVTLHDNNRLTEKKVPINLWIDGKQ---TTVPIDK-VKTSKKEVTVOELDL-QARH--	161
Db	108	CDICTAANS---NKAHPVSNAIDGTERWQSPPLSRGLEYNVNTV---LDLGQVFHVA	160
Qy	162	YLHGKFG-----LYNSDSFGKVGQGLIVHSSRGSTSVSDLDLDAQOQYPTLLRI	212
Db	161	YVLKFNANSPRPDLWVLRSTDFGHTYQWQF-FASSK-----RDCLERFG--PTLERI	212
Qy	213	YRDNTTISSTLS	225
Db	213	TQDDDVICTEYS	225

RESULT 30

DNAC_STRAG	STANDARD;	PRT;	609 AA.
AC	P95693;		
DT	15-JUL-1998 (Rel. 36, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa protein) (HSP70).		
GN	DNAC.		
OS	Streptococcus agalactiae.		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=1311;		
RN	(1)		
RP	SEQUENCE FROM N.A.		
RL	Rioux C.R., Martin D., Hamel J., Brodeur B.R.;		
RL	Submitted (SEP-1996) to the EMBL/genbank/DBJ databases.		
CC	-!- FUNCTION: ACTS AS A CHAPERONE (BY SIMILARITY).		
CC	-!- INDUCTION: BY STRESS CONDITIONS E.G. HEAT SHOCK (BY SIMILARITY).		

```
CC -!- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 70 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U72719; AAB39219.1; -.
DR HSP: P04475; LDG4.
DR InterPro: IPR001023; Hsp70.
DR Pfam: PF00012; HSP70; 1.
DR PRINTS: PR00301; HEATSHOCK70.
DR ProDom: PD000089; Hsp70.1; 1.
DR PROSITE: PS00297; HSP70.1; 1.
DR PROSITE: PS00329; HSP70.2; 1.
DR PROSITE: PS01036; HSP70.3; 1.
KW Chaperone; ATP-binding; Heat shock; Phosphorylation.
KW MOD_RES 173 173 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 609 AA: 64980 MW; 298D8ADCC9D31E0C CRC64;
Query Match 6.7%; Score 81; DB 1; Length 609;
Best Local Similarity 27.0%; Pred. No. 15;
Matches 37; Conservative 19; Mismatches 63; Indels 18; Gaps 5;
QY 69 VDLGST-AATSEYEGSSVDLYGAYGYOCAGGTPNKTACMYGVTLDHNNRLTEKKVPI 127
DB 7 IDLGTNSAVALEGTESKIANPENGRNTPSPVSFKNGEIIIVGDAKQAVTNPDPT 63
QY 128 NLWIDGKQTTVPIDKVTYSKKEVTYQELDLQARHYLHGKFLGNSDFGKVGQRLIVFH 187
DB 64 VISIKSKMGT--SEKVSANGKEVTPQEISAMTLOYLKG----YAEDYLGEKVEKAVI--- 114
QY 188 SSEGSTVSYDLFDAQOQ 204
DB 115 -----TVPAYFNDAQOQ 126
RESULT 31
WDRL_CAEEL
ID WDRL_CAEEL STANDARD; PRT; 611 AA.
AC Q11176;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Actin interacting protein 1 (AIP1) (Uncoordinated protein 78).
GN UNC-78 OR C04F6.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoides;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=21157410; PubMed=11257131;
RA Ono S.;
RT "The Caenorhabditis elegans unc-78 gene encodes a homologue of
RT actin-interacting protein 1 required for organized assembly of muscle
RT actin filaments."
RL J. Cell Biol. 152:1313-1320(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RA Nham M.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Induces disassembly of actin filaments in conjunction
CC with ADF/cofilin family proteins. Regulator of actin organization
CC in myofibrils.
CC -!- SIMILARITY: CONTAINS 9 WD REPEATS (TRP-ASP DOMAINS).
CC -!- SIMILARITY: BELONGS TO THE AIP1 FAMILY OF WD-REPEAT PROTEINS.
CC -----
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF324437; AAK11613.1; -.
DR WormPep: C04F6.4; C503924.
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 9.
DR PRINTS: PR00320; GPROTEINBRPT.
DR ProDom: PD000018; WD40; 2.
DR SMART: SM00320; WD40; 10.
DR PROSITE: PS00678; WD_REPEATS_1; 3.
DR PROSITE: PS50082; WD_REPEATS_2; 6.
DR PROSITE: PS50294; WD_REPEATS_REGION; 1.
KW Actin-binding; Cytoskeleton; Repeat; WD repeat.
FT REPEAT 57 96
FT REPEAT 145 185
FT REPEAT 188 227
FT REPEAT 237 276
FT REPEAT 322 361
FT REPEAT 446 485
FT REPEAT 489 528
FT REPEAT 534 573
FT REPEAT 579 610
FT REPEAT 611 611
SQ SEQUENCE 611 AA: 65323 MW; 821452C661B5D27A CRC64;
Query Match 6.7%; Score 81; DB 1; Length 611;
Best Local Similarity 18.7%; Pred. No. 15;
Matches 53; Conservative 46; Mismatches 78; Indels 106; Gaps 12;
QY 23 LGNLKQIYYNKKATTSSEKSDAQELTNLLPKGFTHG--PW-----YNDLLVDLGT 74
DB 313 LGSIDQVRYGHNKAITALSSAD---GKTLFSADAEGHINSWDISTGTSNRVFPDVHAT 368
QY 75 AATSEYEGSSVDLYGAY-----GYOCAGGTPNK-----TA 105
DB 369 MITGIKTTYSKGLFTVSWDDHLKVVYPAGSGVDDSSKAVANKLSSQPLGLAVSADGDI 428
QY 106 CMYGGVTLHNNRLTEKKVPI-----LWIDGKQTTVPIDKV---KTSKK 148
DB 429 ACYKHAIVSHGKLT-----VPISYNSCVALSNDKQFVAVGDSKVHYVYKLSGASVSE 485
QY 149 EYTVQELDI-----QAR-----HYLHGKFLGNSDFGKVGQRLIVFHS 188
DB 486 KTIIVHPAEITSAFNSNGAFLVATDQSRKVIPIYSVANNFELAHNSW-----TFHT 536
QY 189 SEGSTVSYDLFDAQOQYPTDLLRIYRDNNTTISSTLSISLYLY 231
DB 537 AKVACVSW-----PDNVRLATGSLDNLNVIV 563
RESULT 32
PURA_METJA
ID PURA_METJA STANDARD; PRT; 345 AA.
AC Q57981;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Adenylosuccinate synthetase (EC 6.3.4.4) (IMP--aspartate ligase)
DE (ADSS) (AMPSase).
GN PURA OR MJ0561.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
```

RE MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RL Science 273:1058-1073(1996).
RC !- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE DE NOVO PATHWAY OF PURINE
CC NUCLEOTIDE BIOSYNTHESIS.
CC !- CATALYTIC ACTIVITY: GTP + IMP + L-aspartate = GDP + phosphate +
CC adenylosuccinate.
CC !- PATHWAY: AMP biosynthesis; first committed step.
CC !- SIMILARITY: BELONGS TO THE ADENYLOSUCCINATE SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67505; AAB98554.1; -
DR HSP; P12283; IADE.
DR TIGR; MJ0361; -
DR InterPro; IPR001114; Asucc_synthase.
DR Pfam; PF00709; Adenylosucc_synt; 1.
DR ProDom; PD001188; Asucc_synthase; 1.
DR TIGRFAMs; TIGR00184; purA; 1.
DR PROSITE; PS01266; ADENYLOSUCCIN_SYN_1; 1.
DR PROSITE; PS00513; ADENYLOSUCCIN_SYN_2; FALSE_NEG.
KW Purine biosynthesis; Ligase; GTP-binding; Complete proteome.
FT NP_BIND 18 24 GTP (POTENTIAL).
FT ACT_SITE 144 144 BY SIMILARITY.
SQ SEQUENCE 345 AA; 37820 MW; 9974DBC30D1DEE72 CRC64;

Query Match 6.6%; Score 80; DB 1; Length 345;
Best Local Similarity 22.6%; Pred. No. 9.3;
Matches 44; Conservative 33; Mismatches 60; Indels 58; Gaps 11;
Qy 4 SEEINEK-DLRKKSLOCTALGNLKIYYNSKAITSSKADQFLTWT----- 51
Db 166 SEEVNALLDRGENVLIEGTQ-GTLLSLYYGYTPYVTSKDTTASSPAADVIGTPKVDVI 224
Qy 52 LIFKGFFT---GHPWYNLLVDLGSTAATSEYSGSSVDLYG-----AYYGYOCAGG 99
Db 225 VVFKTFPTRVGAGPPTMSLE-----EAESLGIVEYGTGTCRRRRVGYDFELA-- 274
Qy 100 TPNKTACHYGGVT-----LHDNNRLTEKKVPINLWDGKQTTVPIDKV 143
Db 275 ---RKACRLNGATQIALGLDKYDKECYGVTEYNKLSKAKEFINK-IE-EVTGVPVTII 329
Qy 144 KTSKKEVTVQELDLQ 158
Db 330 STGPE--MHQTIDLR 342

RESULT 33
PCK_METH
ID PKG_METH STANDARD; PRT; 411 AA.
AC O27121;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PKG OR MHI042.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;

OC Methanobacteriaceae; Methanothermobacter.
OX NCBI_TaxID=187420;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Delta H;
RC MEDLINE=98037514; PubMed=9371463;
RX Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,
RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC !- CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC !- PATHWAY: Second phase of glycolysis; second step.
CC !- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC !- SUBCELLULAR LOCATION: Cytoplasmic.
CC !- SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000876; AAB85538.1; -
DR HSP; P36204; 1VPE.
DR InterPro; IPR001576; PKG.
DR Pfam; PF00162; PKG; 1.
DR PRINTS; PR00477; PHGLYCKINASE.
DR PROSITE; PS00111; POLYGLYCERATE_KINASE; 1.
KW Transferase; Kinase; Glycolysis; Complete proteome.
SQ SEQUENCE 411 AA; 44997 MW; E82D9737C6F74F76 CRC64;

Query Match 6.6%; Score 80; DB 1; Length 411;
Best Local Similarity 22.5%; Pred. No. 11;
Matches 47; Conservative 35; Mismatches 75; Indels 52; Gaps 11;
Qy 7 INEKDLRKKSLOCTALGNLKIYYNSKAITSS-----EKSADQFLTWTLLFK 55
Db 184 VMRELR---TLQG-ALENRPCCVYLVGGVKVDDSIMVMKNVLNENGSAIDLVTGLVAN 239
Qy 56 GFFTGHPWYNLLVDLGSTAATSEYSGSSVDLYGYGYOCAGGTPNKTACMYGGVTLLHD 115
Db 240 IFLAG-----CGVKIGKV-----NMDFIKS-RGY-----CDFIKVAKKL 272
Qy 116 NNRUTEKKVPINLWI---DGKQTTVPIDKVKTSK-KVTVVQELDLQARH-----YLHG 165
Db 273 KKRFPERIVWPVDVAVCRDGRKRDVPVKKIPNHPIDQIDGMETIKLYARRIREARTLFG 332
Qy 166 KFLYNSDSFGKVGQGLIVFHSSEGSTV 194
Db 333 PAGVFENPDFSIGTEDILNLAISSEGSFSI 361

RESULT 34
ITAL_HUMAN
ID ITAL_HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=9315124; PubMed=8428973;
RX Briesewitz R., Epstein M.R., Marcantonio E.E.;
RT "Expression of native and truncated forms of the human integrin alpha
1 subunit";
RL J. Biol. Chem. 268:2989-2996(1993).
CC -!- FUNCTION: INTEGRIN ALPHA-1/BETA-1 IS A RECEPTOR FOR LAMININ AND
COLLAGEN. IT RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-F-P-G-
CC E-R IN COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-1
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VFMA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: BELONGS TO THE INTEGRIN ALPHA CHAIN FAMILY.
CC -!- SIMILARITY: CONTAINS 1 VFMA DOMAIN.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC -!- SIMILARITY: CONTAINS 7 FG-GAP REPEATS.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD49a entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd49a.htm".
DR HSSP: p17301; LAOX.
DR Genew: HGNC:6134; ITGAL.
DR MIM: 192968;
DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF00092; vwa; 1.
DR Pfam: PF00357; Integrin_A; 1.
DR Pfam: PF01839; FG-GAP; 5.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00191; Int_alpha; 5.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS0234; VWF_A; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Repeat; Calcium; Magnesium.
FT DOMAIN 1 1113 EXTRACELLULAR (POTENTIAL).
FT TRANSMPM 1114 1136 POTENTIAL.
FT DOMAIN 1137 1151 CYTOPLASMIC (POTENTIAL).
FT REPEAT 16 75 FG-GAP 1.
FT REPEAT ? 2 FG-GAP 2.
FT DOMAIN 147 360 VWA.
FT REPEAT 349 404 FG-GAP 3.
FT REPEAT 405 457 FG-GAP 4.
FT REPEAT 459 520 FG-GAP 5.
FT REPEAT 540 599 FG-GAP 6.
FT REPEAT 602 654 FG-GAP 7.
FT CA_BIND 470 478 POTENTIAL.
FT CA_BIND 552 560 POTENTIAL.
FT CA_BIND 614 622 POTENTIAL.
FT SITE 1139 1142 GFFRR MOTIF.
FT DISULFID 54 64 BY SIMILARITY.
FT DISULFID 660 669 BY SIMILARITY.
FT DISULFID 675 728 BY SIMILARITY.
FT DISULFID 780 786 BY SIMILARITY.
FT DISULFID 850 858 BY SIMILARITY.
FT DISULFID 1002 1034 BY SIMILARITY.
FT DISULFID 1037 1044 BY SIMILARITY.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 84 84 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 374 374 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 432 432 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 504 504 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 720 720 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 752 752 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 855 855 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 887 887 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 911 911 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 938 938 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 946 946 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 980 980 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1055 1055 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1151 AA; 127837 MW; 6B3F3C1AABF52808 CRC64;
Query Match 6.6%; Score 80; DB 1; Length 1151;
Best Local Similarity 27.3%; Pred. No. 41;
Matches 53; Conservative 25; Mismatches 66; Indels 50; Gaps 12;
Qy 72 GSTAATSEYEGSSV-----DLYGAYGYOCAGTGNKTKCMYGGVTL----- 113
Db 586 GGDGKTLKFFGQSIHGEMDLNG-----DGLTDVTITGGLGGAALFWSRDVAVVKTMTN 647
Qy 114 HDNNRLTEEEKVPINLWIDGKOT-----TVPIDKVKTSKKEVTVOELDLQARHYLHGKFG 168
Db 648 FEPKNVNIQKK---NCHWEGKETVCINATVCFE-VLKSKEDTIYEAOLQYRTLDLSLRQ 703
Qy 169 LYNDSDFGG---KYQGLIVFHSSEGSTSVSYDLFDAQGOYPTDLLRIYRD-----NT 217
Db 704 ISRS-FFSGTQERKQVQRTIV-RKSECTKHSFYMLD-KHDFODS-VRITLDNLNLTDPENG 759
Qy 218 TISSTSLISLYLY 231
Db 760 PVLDDSLPNSVHEY 773
RESULT 35
YE64_PASMU
ID YE64_PASMU STANDARD; PRT; 324 AA.
AC Q9CKY6;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical paba-like protein PMI464.
GN PMI464.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -!- SIMILARITY: TO THE C-TERMINAL OF PARA-AMINOBENZOATE SYNTHASE
CC COMPONENT 1.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006183; AAK03548.1; -
CC InterPro; IPR000350; Chorisimate_bind.
CC Pfam; PF00425; chorisimate_bind; 1.
CC PRINTS; PR00095; ANTSNTHASE1.
CC PRODOM; PD000779; Chorisimate_bind; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 324 AA; 37309 MW; E9336FF3890A32AE CRC64;

```
Query Match          6.5%; Score 79.5; DB 1; Length 324;
Best Local Similarity 27.3%; Pred. No. 9.5;
Matches 24; Conservative 17; Mismatches 34; Indels 13; Gaps 3;

QY 136 TTVPIDKVKTSKKEVTYQEL---DLQARHYLHGKFGLYNSDSFGKGVORGLI-----V 185
      | : | : | | | : | : | | | : | | | : | | : | | : | | : | | : |
DB 236 TLLPAGSISGAPKKTQVGIHAAEQRPGRYTGIFGLGDSGLQSAVAIRTEQVDEKLI 295
      | : | : | | | : | : | : | : | : | : | : | : | : | : | : | : |
QY 186 FHSSEGSTVSVYDLFDAQOQYPTLLRIY 213
      | : | : | | | : | : | : | : | : | : | : | : | : | : | : | : |
DB 296 FRSGGGITILSELED---EYQELIQKV 320
      | : | : | | | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 36
VEL_HPV35
ID VEL_HPV35 STANDARD; PRT; 637 AA.
AC
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Replication protein E1.
GN E1.
OS Human papillomavirus type 35.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10587;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94265501; PubMed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92124753; PubMed=1310198;
RA Marich J.E., Pontstler A.V., Rice S.M., McGraw K.A., Dubensky T.W.;
RT "The phylogenetic relationship and complete nucleotide sequence of
human papillomavirus type 35.";
RL Virology 186:770-776(1992).
CC
CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE. REQUIRED FOR INITIATION OF
VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; X74477; CRA52553.1; -.
CC EMBL; M74117; AAA46968.1; -.
CC PIR; A40824; W1WL35.
CC PIR; S36523; S36523.
CC InterPro: IPR001177; Papillom_E1.
CC Pfam; PF00519; El; 1.
CC Pfam; PF00524; ELN; 1.
KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW Nuclear protein.
FT NP_BIND 463 470 ATP (POTENTIAL).
FT CONFLICT 31 34 DPVS -> SSV (IN REF. 2).
FT CONFLICT 142 143 QQ -> HE (IN REF. 2).
FT CONFLICT 235 258 ESKTLKPYCLYIHQICLSWG ->
NFKHTYVYINVYRVHGA (IN REF. 2).
FT CONFLICT 269 281 CAKNRTTIEKLS -> VERREQQLTKIDA (IN REF.
2).
FT CONFLICT 417 418 EK -> AQ (IN REF. 2).
FT CONFLICT 515 552 WAYIDQYLRNALDGNPISLDVKKALVOLKCPPLLINS ->
GIYPIEFKCTRWKYSISFRCKALSIHIMPTFTYII (IN
REF. 2).

FT CONFLICT 587 587 V -> E (IN REF. 2).
FT CONFLICT 613 613 D -> V (IN REF. 2).
SQ SEQUENCE 637 AA; 72122 MW; 3C694D4451791003 CRC64;

Query Match          6.5%; Score 79.5; DB 1; Length 637;
Best Local Similarity 20.3%; Pred. No. 22;
Matches 50; Conservative 39; Mismatches 88; Indels 69; Gaps 11;

QY 21 TALGNLKIYYIYNKAITSSSEKSADOFLLTNTLLPKGFTGHPWYNDLLVLDLGSNTAATSEY 80
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 310 TAMSNISEV-----DGETPEWIORQTVLQHSF-----NDAIFDL-SEMVQWAY 351
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 81 EGSSVDLYGAYGYQCAGGTGNKTCAMY-----GGVTLHNNRLTEEKVPINL 129
      | : | : | | | | | | | | | | | | | | | | | : | : | : | : |
DB 352 DNDFIDSDSIAYKYAQLAET-NSNACAPLKSNSQAKIVKDCATMCRHYKRAKREMTNSQ 410
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 130 WIDGKQTTVPID-KVKTSKKEVTYQELDLQA-----RHYLHGK-----FGLYNSDSF 175
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 411 WIKRRCEKVDGDDGWRDIVRFRLYQQVDFVAFLSALKNFLHGVPKKNKILYGAAPT--- 467
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 176 GKVQVQGLIVFHSSEGSTVSY-----DLFDAQ-----GQYPTDLLRIYR 214
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 468 -GKSLFGSLMHFLQGLAISYVNSKSHFWLQPLDYDAKTAMLDDATSPCWAYIDQYLRNAL 526
      | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 215 DNTTIS 220
      | |
DB 527 DGNPIS 532

RESULT 37
Y650_METJA
ID Y650_METJA STANDARD; PRT; 692 AA.
AC Q57852;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MJ0650 precursor.
GN MJ0650.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RL Science 273:1058-1073(1996).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
CC EMBL; U67512; AAB98647.1; -.
CC TIGR; MJ0650; -.
KW Hypothetical protein; Signal; Complete proteome.
FT SIGNAL 1 24 POTENTIAL
FT CHAIN 25 692 HYPOTHETICAL PROTEIN MJ0650.
SQ SEQUENCE 692 AA; 80607 MW; 25D0A91A08177188 CRC64;
```

```

DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hydrolase; Glycosidase; Chitin degradation; Signal; Glycoprotein;
KW Endoplasmic reticulum.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 551 PROBABLE ENDOCHITINASE.
FT ACT_SITE 305 305 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 173 173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 548 551 PREVENT SECRETION FROM ER (POTENTIAL).
SQ SEQUENCE 551 AA; 61368 MW; 4DDAAD187873BBA2 CRC64;

Query Match 6.5%; Score 79; DB 1; Length 551;
Best Local Similarity 23.8%; Pred. No. 20;
Matches 44; Conservative 22; Mismatches 71; Indels 48; Gaps 11;

QY 32 YNSKAITSS---SEKSADQFLNTLLEKGFHGWYNDLLVDLGSTAATSEVEGSSVDLY 88
DB 358 YDKIAVNVYAPQAQSLKGIIFLMSYDFKG-----AWSN---TDLGVTQTVVYAPSWNSEELY 409
QY 89 GAYGYQC---AGGTPNKNT---ACMYG---GVTLHDN-NRLTEKKVP-INLWIDGKQT 136
DB 410 TTHAVDALLKQGVDPNKIIIVGVAMYGRGWTGVNTYNDVYFGTNGNGPGSGTWEDG--- 466
QY 137 TVPTDKVTSKKEVTYQELDQARHYLHGKGLYNSDSFGKGKVGORGLIVFHSSEGSTVSY 196
DB 467 -----VVDYRQIQKDLNNY-----VYTFDS-----AAQASYVDFKSGKDLISF 504

QY 197 DLFDA 201
DB 505 DSVDS 509

RESULT 39
SLA2_BACAN
ID SLA2_BACAN STANDARD; PRT; 862 AA.
AC P94217;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S-layer protein EAL precursor.
ES EAG.
GN Bacillus anthracis.
OS Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OC NCBI_TaxId=1392;
OX [1]
RN RC
RP SEQUENCE FROM N A.
RC STRAIN=Sterne / 9131;
RX MEDLINE=37260111; PubMed=9106206;
RA Mesnage S., Tosi-Couture E., Mock M., Gounon P., Fouet A.;
RT "Molecular characterization of the Bacillus anthracis main S-layer
RT component: evidence that it is the major cell-associated antigen.";
RL Mol. Microbiol. 23:1147-1155(1997).
CC -1- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC -1- SUBCELLULAR LOCATION: Cell wall.
CC -1- SIMILARITY: CONTAINS 3 S-LAYER HOMOMOLOGY (SLH) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X99724; CAA68063.1; -.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00395; SLH; 3.
KW Signal; Repeat; Cell wall; S-layer.
FT SIGNAL 1 29 POTENTIAL.
FT CHAIN 30 862 S-LAYER PROTEIN EAL.
FT DOMAIN 34 76 SLH 1.
FT DOMAIN 95 136 SLH 2.

```


DR InterPro; IPR000034; Laminin_B.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR001791; Laminin_G.
DR Pfam; PF00053; laminin_EGF; 2.
DR Pfam; PF00054; laminin_G; 2.
DR ProDom; PD002082; Laminin; 1.
DR ProDom; PD003031; Laminin_B; 1.
DR ProSITE; PS00022; EGF_1; 19.
DR ProSITE; PS01186; EGF_2; 3.
DR ProSITE; PS01248; LAMININ_TYPE_EGF; 19.
DR ProSITE; PS50025; LAM_G_DOMAIN; 5.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
FT SIGNAL 1 35
FT CHAIN 36 3695
FT DOMAIN 36 299 LAMININ ALPHA-5 CHAIN.
FT DOMAIN 300 358 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 359 428 LAMININ EGF-LIKE 1.
FT DOMAIN 429 474 LAMININ EGF-LIKE 2.
FT DOMAIN 494 540 LAMININ EGF-LIKE 3.
FT DOMAIN 541 586 LAMININ EGF-LIKE 4.
FT DOMAIN 587 631 LAMININ EGF-LIKE 5.
FT DOMAIN 632 676 LAMININ EGF-LIKE 6.
FT DOMAIN 677 722 LAMININ EGF-LIKE 7.
FT DOMAIN 723 775 LAMININ EGF-LIKE 8.
FT DOMAIN 776 828 LAMININ EGF-LIKE 9.
FT DOMAIN 829 850 LAMININ EGF-LIKE 10.
FT DOMAIN 851 1437 LAMININ EGF-LIKE 11 (INCOMPLETE).
FT DOMAIN 1438 1483 LAMININ DOMAIN IV 1 (DOMAIN IV B).
FT DOMAIN 1484 1527 LAMININ EGF-LIKE 12.
FT DOMAIN 1528 1577 LAMININ EGF-LIKE 13.
FT DOMAIN 1578 1627 LAMININ EGF-LIKE 14.
FT DOMAIN 1628 1637 LAMININ EGF-LIKE 15.
FT DOMAIN 1638 1830 LAMININ EGF-LIKE 16 (N-TERMINAL).
FT DOMAIN 1831 1863 LAMININ EGF-LIKE 17.
FT DOMAIN 1864 1912 LAMININ EGF-LIKE 18.
FT DOMAIN 1913 1968 LAMININ EGF-LIKE 19.
FT DOMAIN 1969 2022 LAMININ EGF-LIKE 20.
FT DOMAIN 2023 2069 LAMININ EGF-LIKE 21.
FT DOMAIN 2070 2116 LAMININ EGF-LIKE 22.
FT DOMAIN 2117 2166 LAMININ EGF-LIKE 23.
FT DOMAIN 2167 2735 DOMAIN II AND I.
FT DOMAIN 2736 2929 LAMININ G-LIKE 1.
FT DOMAIN 2930 3115 LAMININ G-LIKE 2.
FT DOMAIN 3116 3292 LAMININ G-LIKE 3.
FT DOMAIN 3293 3340 LAMININ G-LIKE 4.
FT DOMAIN 3341 3513 LAMININ G-LIKE 5.
FT DOMAIN 3514 3692 LAMININ G-LIKE 6.
FT DOMAIN 3693 4000 COILED COIL (POTENTIAL).
FT DOMAIN 4001 4221 COILED COIL (POTENTIAL).
FT DOMAIN 4222 4400 COILED COIL (POTENTIAL).
FT DOMAIN 4401 4588 CELL ATTACHMENT SITE (POTENTIAL).
FT DOMAIN 4589 4724 CELL ATTACHMENT SITE (POTENTIAL).
FT DOMAIN 4725 4840 CELL ATTACHMENT SITE (POTENTIAL).
FT DOMAIN 4841 5000 BY SIMILARITY.
FT DOMAIN 5001 5150 BY SIMILARITY.
FT DOMAIN 5151 5266 BY SIMILARITY.
FT DOMAIN 5267 5338 BY SIMILARITY.
FT DOMAIN 5339 5415 BY SIMILARITY.
FT DOMAIN 5416 5500 BY SIMILARITY.
FT DOMAIN 5501 5560 BY SIMILARITY.
FT DOMAIN 5561 5627 BY SIMILARITY.
FT DOMAIN 5628 5710 BY SIMILARITY.
FT DOMAIN 5711 5840 BY SIMILARITY.
FT DOMAIN 5841 5999 BY SIMILARITY.

FT DISULFID 589 605 BY SIMILARITY.
FT DISULFID 607 616 BY SIMILARITY.
FT DISULFID 619 629 BY SIMILARITY.
FT DISULFID 632 644 BY SIMILARITY.
FT DISULFID 634 650 BY SIMILARITY.
FT DISULFID 652 661 BY SIMILARITY.
FT DISULFID 664 674 BY SIMILARITY.
FT DISULFID 677 689 BY SIMILARITY.
FT DISULFID 679 696 BY SIMILARITY.
FT DISULFID 698 707 BY SIMILARITY.
FT DISULFID 710 720 BY SIMILARITY.
Query Match Score 79; DB 1; Length 3695;
Best Local Similarity 29.4%; Pred. No. 2.1e-02;
Matches 58; Conservative 20; Mismatches 75; Indels 44; Gaps 13;
QY 56 GFTGHPWYNDLLVDLSTAAATSEYEGSSV-----DLYGAYGYQCAGGTNKT- 104
DB 41 GGFSLHPYPYFNLAEE--GARIAASATCGEAPARGSPRTEDLYCKLVGGPVAGGDPNQTI 98
QY 105 ACMYGGVTLHDNNRLTEKKVPINLWIDGKQ---TTVPIDK-VKTSKKEVTVOELDL-QA 159
DB 99 OGQYCDICTAANS-----NKAHPASNAIDGTERWQSPPLSRGLEYNVWVT---LDLQGV 151
QY 160 RH--YUHGKFG-----LYNSDSFGKVGQRLVIFHSSEGSTVSYDLFDAGQGYPDT 208
DB 152 FHVAVYLIKFNANSPRDLWLERSMDFGRTYQWQF-FASSK-----RDLERFG--PQT 203
QY 209 LLRIYRDNTTISSTLS 225
DB 204 LERITRDAACITTEYS 220
RESULT 41
YJRL_YEAST
ID YJRL_YEAST STANDARD; PRT; 396 AA.
AC P46992;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Hypothetical 43.0 kDa protein in CPSI-FPPI intergenic region.
GN YJL171C OR J0512.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OX Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rinke M., Domdey H.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO YEAST YBR162C.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; 249446; CAAB9466.1; -.
DR SGD; S0003707; YJL171C.
KW Hypothetical protein.
SQ SEQUENCE 396 AA; 43014 MW; 279E858E7512670A CRC64;
Query Match Score 78.5; DB 1; Length 396;
Best Local Similarity 21.6%; Pred. No. 15;
Matches 64; Conservative 29; Mismatches 86; Indels 117; Gaps 16;
QY 26 LKQIYYNSKAIT-----SSEKSDQF-----LNTTLF-K 55
DB 85 LKQAFYTPGFTVNNSSSDNNRLAYESSKTDADNVTFLNHGGEASPCGLNALSAYS 144

```
QY 56 GFTTCHPWYNLLVDLGGTAATSEY-----EGSSVDLYG--AYYGVQCAGG 99
DB 145 SNGTGSASEATVLADGTEYLIIISNVSCPKSGYDKGCVYRSRGPAYGY---GG 201
QY 100 T-----PNKTACMYGGVTLHD-----NNRLTEKKVVPIN-----LWIDGKQTTV 138
DB 202 TTKMFLFEFEMTEPEKNSSSIGYDLPAILNLDHIARTSOYPTNANCSWASG-----256
QY 139 PIDKVKTSKKEVTQVQELD-----QARHLYHGKGLYNS-DSFGGKVO-----180
DB 257 -----CCEYDIFEMNCTENKHLVSPHTFGQIEDLGTGQISQVGYTRNTG 303
QY 181 --RGLIVFHSSEGVSVYDLFDA-----QCGYPTDILLRIYRDNTTSSLSLS 227
DB 304 TMKGGVWFDS--GNVVF--ISDAPFNGTVSADTVNDLLAIPENETYSQQLMSIS 357

RESULT 42
YGS0_YEAST STANDARD; PRT; 413 AA.
AC P45819;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE Hypothetical 47.7 kDa protein in ROK1-SUA5 intergenic region.
GN YGL170C OR G1654.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / FV1679;
RX MEDLINE=97051590; PubMed=8896267;
RA Klima R., Coglievina M., Zaccaria P., Bertani I., Bruschi C.V.;
RT "A putative helicase, the SUA5, PMR1, tRNAlysl genes and four open
RT reading frames have been detected in the DNA sequence of an 8.8 kb
RT fragment of the left arm of chromosome VII of Saccharomyces
RT cerevisiae."
RL Yeast 12:1033-1040(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X85757; CAA59759.1; -
DR EMBL; Z72692; CAA96882.1; -
DR SGD; S0003138; YGL170C.
KW Hypothetical protein.
SQ SEQUENCE 413 AA; 47700 MW; ACEB8DA9A240B650 CRC64;

Query Match 6.4%; Score 78.5; DB 1; Length 413;
Best Local Similarity 20.3%; Pred. No. 16;
Matches 47; Conservative 37; Mismatches 89; Indels 59; Gaps 9;

QY 1 SEKSEINEKDLRKSELOGTALGNLKOIYYNSKATSS-----41
DB 41 SKRNSPVAKQS--QKDEEKSMMGTASIFHENKDIHERSEHTDDFNDGLKLAPDSSPSL 98
QY 42 -----KSADQFLNTLLFKGFTTGH--PWYNDLLVDLGTSTAATSEYSGSSVDLYGAYGY 94
DB 99 KECQFKNWEFWCNT---EGYTKMKQFPH-----FTSGLEIKPVPVNEINISTSPY 147
QY 95 QCAGGTPNKACMYGGVTLHDNNRLTEKKVVPINLW-----IDGKQTTVPi--DKVKTS 146
DB 148 K--GQRNPSAPTEYSAAATTAFTKTOLEVSFLKTNLLTVIKKEIDICLSVFFDFAVQMQ 205
QY 147 KKEVTQVQLDQARHLYHGKGLYNSDSFGGKVGRLVIFHSSSGSTVSYDL 198
DB 147 KKEVTQVQLDQARHLYHGKGLYNSDSFGGKVGRLVIFHSSSGSTVSYDL 198
```

```
DB 206 KKFLEYRIDLDDEEVEL-----KILGELLNDLNFHMQENSLNREL 247

RESULT 43
SLAP_ACEKI STANDARD; PRT; 762 AA.
ID SLAP_ACEKI 22258;
AC 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cell surface protein precursor (S-layer protein).
OS Acetogenium kivui.
OC Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC Thermoanaerobacteriaceae; Thermoanaerobacter.
OX NCBI_TaxID=2325;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=DSM 2030;
RX MEDLINE=90036724; PubMed=2681162;
RA Peters J., Peters M., Lottspeich F., Baumeister W.;
RT "S-layer protein gene of Acetogenium kivui: cloning and expression in
RT Escherichia coli and determination of the nucleotide sequence."
RL J. Bacteriol. 171:6307-6315(1989).
RN [2]
RP PARTIAL SEQUENCE, AND CARBOHYDRATE-LINKAGE SITES.
RC STRAIN=DSM 2030;
RX MEDLINE=92281680; PubMed=1596358;
RA Peters J., Rudolf S., Oschkinat H., Mengele R., Sumper M.,
RA Kellermann J., Lottspeich F., Baumeister W.;
RT "Evidence for tyrosine-linked glycosaminoglycan in a bacterial
RT surface protein."
RL Biol. Chem. Hoppe-Seyler 373:171-176(1992).
RN [3]
RP DOMAINS.
RX MEDLINE=94156823; PubMed=8113161;
RA Lupas A., Engelhardt H., Peters J., Santarius U., Volker S.,
RA Baumeister W.;
RT "Domain structure of the Acetogenium kivui surface layer revealed by
RT electron crystallography and sequence analysis."
RL J. Bacteriol. 176:1224-1233(1994).
CC -!- FUNCTION: THE S-LAYER IS A PARACRYSTALLINE MONO-LAYERED ASSEMBLY
CC OF PROTEINS WHICH COAT THE SURFACE OF BACTERIA.
CC -!- SUBCELLULAR LOCATION: CELL WALL. THIS BACTERIUM IS COVERED BY A
CC S-LAYER WITH HEXAGONAL SYMMETRY.
CC -!- PTM: THE CARBOHYDRATE CONTENT OF THIS PROTEIN IS ABOUT 8% WHICH
CC CORRESPOND TO ABOUT 40 TO 50 SUGAR MOLECULES PER MONOMER. O-LINKED
CC GLYCANS CONSIST OF GLC, GALNAc AND GLCNAC.
CC -!- SIMILARITY: CONTAINS 3 S-LAYER HOMOLGY (SLH) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31069; AAA21930.1; -
DR PIR; A34355; A34355.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00395; SLH; 2.
DR PROSITE; PS01072; SLH_DOMAIN; 2.
KW Signal; Glycoprotein; Repeat; Cell wall; S-layer.
FT SIGNAL 1 26
FT CHAIN 27 762 CELL SURFACE PROTEIN.
FT DOMAIN 30 94 SLH 1.
FT DOMAIN 95 145 SLH 2.
FT DOMAIN 146 204 SLH 3.
FT DOMAIN 473 479 SER/THR-RICH.
FT DOMAIN 625 630 SER/THR-RICH.
FT CARBOHYD 297 297 O-LINKED (GLC. . .)
FT CARBOHYD 516 516 O-LINKED (GLC. . .)
FT CARBOHYD 520 520 O-LINKED (GLC. . .)
```

```
FT CARBOHYD 632 632 O-LINKED (GLC...).
SQ SEQUENCE 762 AA; 82785 MW; 34EC9C784DECA67E CRC64;

Query Match
Best Local Similarity 6.4%; Score 78.5; DB 1; Length 762;
Matches 60; Conservative 28; Mismatches 105; Indels 49; Gaps 11;

QY 3 KSEINEKDLRKKSELOQTALGNLKOIYYY-----NSKAITSEKSAQDQFTLTLFK 55
DB 90 KSEKSAFQVDPQNH---WAVGQINLAYKLGIAQGVGNCKFDPNSRLRYAALAFVLRAL 145
QY 56 GFETGHPWNLVDLGGTAATSEYEGSSVDLYGAYGYQCAGGTPNKATACWYGGVTLHD 115
DB 146 GF-----KDLDPYG-----YLAQAQDL-GLVHGLNLA-----YNGVIKRG 180
QY 116 NNRLTTEK--KVPINLWIDGKOTTP--IDKVKTSKKEVTVOELDLQARHYLHGKFLYN 171
DB 181 DLALILDLRALEVPWKYVDGKVEGLKSLKAT--KAETVTIATNAQDSVEEGKVAVLD 239
QY 172 SDSPGKVGQRLIVFHSSEGSTVSVDLFDAAQOQYPTDLLRIYRDNNTISST----SLSIS 227
DB 240 KDGKLTINAGLVDFSEVLGKKV----IVYSERFGDPVYVAEGDNDVVSFTTQSDSVGTT 295
QY 228 LY 229
DB 296 VY 297

RESULT 44
ADH1_ORYSA
ID ADH1_ORYSA STANDARD; PRT; 376 AA.
AC P20306;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Alcohol dehydrogenase 1 (EC 1.1.1.1).
GN ADH1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Indica-IR26; TISSUE=Seedling;
RX MEDLINE=93357433; PubMed=2562760;
RA Xie Y., Wu R.;
RT "Rice alcohol dehydrogenase genes: anaerobic induction, organ
specific expression and characterization of cDNA clones.";
RL Plant Mol. Biol. 13:53-68(1989).
CC -|- CATALYTIC ACTIVITY: An alcohol + NAD(+) -> an aldehyde or ketone +
NADH.
CC -|- COFACTOR: ZINC.
CC -|- SUBUNIT: HOMODIMER.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X16296; CAA34363.1; -
DR PIR; JQ0474; JQ0474.
DR HSSP; P11766; ITEX.
DR InterPro; IPR002328; ADH_zinc.
DR InterPro; IPR002085; Adh_zn_family.
DR InterPro; IPR000205; NAD_binding.
DR Pfam; PF00107; adh_zinc; 1.
DR PROSITE; PS00059; ADH_ZINC; 1.
KW Oxidoreductase; Zinc; NAD; Multigene family.
```

```
FT METAL 45 45 ZINC (CATALYTIC).
FT METAL 67 67 ZINC (CATALYTIC).
FT METAL 98 98 ZINC (SECOND ATOM).
FT METAL 100 100 ZINC (SECOND ATOM).
FT METAL 103 103 ZINC (SECOND ATOM).
FT METAL 111 111 ZINC (SECOND ATOM).
FT METAL 175 175 ZINC (CATALYTIC).
SQ SEQUENCE 376 AA; 40852 MW; B8ED6E025863D64F CRC64;

Query Match
Best Local Similarity 6.4%; Score 78; DB 1; Length 376;
Matches 47; Conservative 29; Mismatches 77; Indels 46; Gaps 7;

QY 3 KSEINEKDLRKKSELOQTALGNLKOIYYNSKAI-----TSSE-----KS 43
DB 104 KSAESNMCDLLINIDRGVMIGDGKSRFSINGKPIYHFGTSTFSEYTVMHVGCVAKINP 163
QY 44 ADQFTLNTLLFGFTGHPWYNDLLVDLGSTAAATSEYEGSSVDLYG-AYYQCAGGTPN 102
DB 164 AAPLDKVCVLSGGISTG-----LGATINVAKPKGSTVAIFGLGAVGLAAABEGARI 213
QY 103 KTACWYGGVTLHDNRLTEKKVPINLWIDGKQTPVPIDKVKTSKKEVTVOELD----- 156
DB 214 RGASRIIGIDL-NANRFEEARKFGCTEFYNPKDHPKQVOOVL-----EMTNGVDRSVECT 269
QY 157 -----LQARHYLHGKFL 169
DB 270 GNINAMIQAFECVHDGWI 288

RESULT 45
DNAK_STRPN
ID DNAK_STRPN STANDARD; PRT; 607 AA.
AC P95829; O66035;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chaperone protein dnaK (Heat shock protein 70) (Heat shock 70 kDa
protein) (HSP70).
GN DNAK OR SP0517.
OS Streptococcus pneumoniae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1313;
RN [1]
RP SEQUENCE FROM N.A.
RA Rioux C.R., Martin D., Hamel J., Brodeur B.R.;
RT "Heat shock protein HSP70 and amino terminus of DnaJ of Streptococcus
pneumoniae.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Rx / CPI200;
RX MEDLINE=98231633; PubMed=9570114;
RA Kim S.-W., Choi I.-H., Kim S.-N., Kim Y.-H., Pyo S.-N., Rhee D.-K.;
RT "Molecular cloning, expression, and characterization of dnaK in
Streptococcus pneumoniae.";
RL FEMS Microbiol. Lett. 161:217-224(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=TIGR4;
RX MEDLINE=21357209; PubMed=11463916;
RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
RA Peterson S., Heidelberg J., DeBoy R.T., Haft D.H., Dodson R.J.,
RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
RA Holtzapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
RA McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,
RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,
RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
RT "Complete genome sequence of a virulent isolate of Streptococcus
pneumoniae.";
RL Science 293:498-506(2001).
```


RP SEQUENCE FROM N.A., AND SEQUENCE OF 34-41.
RC STRAIN-B6A-RI;
RX MEDLINE=94028940; PubMed=8215382;
RA Lee Y.-E., Lowe S.E., Zeikus J.G.;
RT "Gene cloning, sequencing, and biochemical characterization of
RT endoxylanase from Thermomicrobacterium saccharolyticum B6A-RI.";
RL Appl. Environ. Microbiol. 59:3134-3137(1993).
RN [2]
RP ACTIVE SITE, AND MUTAGENESIS.
RX MEDLINE=93388520; PubMed=8376336;
RA Lee Y.-E., Lowe S.E., Henrissat B., Zeikus J.G.;
RT "Characterization of the active site and thermostability regions of
RT endoxylanase from Thermomicrobacterium saccharolyticum B6A-RI.";
RL J. Bacteriol. 173:5890-5898(1993).
RN [2]
CC -!- FUNCTION: ENDO-ACTING ENZYME THAT RANDOMLY CLEAVES THE INTERNAL
CC XYLIDISACCHARIDES OF THE XYLAN BACKBONE, YIELDING
CC XYLIDISACCHARIDES OF VARIOUS LENGTHS, WHICH ARE FURTHER
CC HYDROLYSED TO XLOSE MOLECULES BY BETA-XYLOSIDASE (EC 3.2.1.37).
CC THE TEMPERATURE AND PH OPTIMUM FOR THIS ENZYME ARE 70 DEGREES
CC CELSIUS AND 5.5, RESPECTIVELY. REQUIRES AT LEAST THREE XLOSE
CC RESIDUES FOR CATALYTIC ACTIVITY. DOES NOT HAVE ACTIVITY AGAINST
CC XYLULOSE.
CC -!- CATALYTIC ACTIVITY: Endohydrolysis of 1,4-beta-D-xylosidic
CC linkages in xylans.
CC -!- PATHWAY: Xylan degradation.
CC -!- INDUCTION: BY XYLAN AND XLOSE.
CC -!- SIMILARITY: BELONGS TO CELLULASE FAMILY F (FAMILY 10 OF GLYCOSYL
CC HYDROLASES).
CC -!- SIMILARITY: CONTAINS 2 S-LAYER HOMOMOLOGY (SLH) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M97882; AAA21812.1; ALT_SEQ.
DR HSP; P10478; 1XVZ.
DR InterPro; IPR003305; CBM_Cenc.
DR InterPro; IPR001547; GH_5.
DR InterPro; IPR001000; Glyco_hydro_10.
DR InterPro; IPR001119; SLH.
DR Pfam; PF00331; Glyco_hydro_10; 1.
DR Pfam; PF00395; SLH; 2.
DR Pfam; PF02018; CBM_4_9; 2.
DR PRINTS; PR00134; GLYDRLASE10.
DR PROSITE; PS00591; GLYCOSYL_HYDROL_F10; 1.
DR PROSITE; PS00659; GLYCOSYL_HYDROL_F5; 1.
DR PROSITE; PS01072; SLH_DOMAIN; 2.
DR Xylan degradation; Hydrolase; Glycosidase; Signal; Repeat.
KW SIGNAL 1 33
FT CHAIN 34 1157 ENDO-1,4-BETA-XYLANASE A.
FT DOMAIN 413 649 CATALYTIC (POTENTIAL).
FT DOMAIN 1055 1113 SLH 1.
FT DOMAIN 1114 1157 SLH 2.
FT ACT_SITE 495 495 PROTON DONOR (BY SIMILARITY).
FT ACT_SITE 537 537
FT ACT_SITE 600 600 NUCLEOPHILE.
FT MUTAGEN 537 537 D->N: LOSS OF ACTIVITY.
FT MUTAGEN 600 600 E->Q: LOSS OF ACTIVITY.
FT MUTAGEN 602 602 D->N: LOSS OF ACTIVITY.
SQ SEQUENCE 1157 AA; 128379 MW; 51FA6004497EC58B CRC64;
Query Match 6.4%; Score 78; DB 1; Length 1157;
Best Local Similarity 17.3%; Pred. No. 62;
Matches 38; Conservative 50; Mismatches 80; Indels 52; Gaps 11;
QY 23 LGNLKQIYYNSKATSSPKSADQFLNTLLPKFGFTGHPWNLNLDVLDGSGTAATSEYE- 81
DB 160 IGTGLKTLWY-----VESPDPTLEYIYDVVV-----TTQNP-----IQGNVIANETFFEN 205

QY 82 -----GSSVDLYGAYGYQCAGG-----TPNKTACMYGCVTLHDNNRLTEKKVPINLW 130
DB 206 GNTSGWIGTSSWRAVYGAHSGDYSLTTGRTA-NNNGPSYDLTGKIVPGQYNVDFW 364
QY 131 ---IDGKQTTVPIDKVKTSKKEVTQELDLQARHY-----LHGKFGLYNSDSFGG 177
DB 265 VKFVNGNDT-----EQIKATVVKATSKDNTYIQVNDNFANVKNKGWTEIKGSFTLPVAD----- 316
QY 178 KVQRGGLIVFHSSEGSVSY--DLFDAQGOYPTDILLRIYRD 215
DB 317 --YSGISIVVESQNTPLEYIIDFVIGISNNQITQND 354
RESULT 48
GTF2_STRDO
ID GTF2_STRDO STANDARD; PRT; 1592 AA.
AC P27470;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Glucosyltransferase-I precursor (EC 2.4.1.5) (GTF-I) (Dextranucrase)
DE (Sucrose 6-glucosyltransferase).
OS Streptococcus downei (Streptococcus sobrinus).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=13117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6715;
RX MEDLINE=91123227; PubMed=1704006;
RA Abo H., Matsumura T., Ohta H., Fukui K., Kato K.,
RA Kagawa H.;
RT "Peptide sequences for sucrose splitting and glucan binding within
RT Streptococcus sobrinus glucosyltransferase (water-insoluble glucan
RT synthetase).";
RL J. Bacteriol. 173:989-996(1991).
CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT
CC TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE
CC OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE
CC AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.
CC -!- CATALYTIC ACTIVITY: Sucrose + {(1,6)-alpha-D-glucosyl}(N) = D-
CC fructose + {(1,6)-alpha-D-glucosyl}(N+1).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DISEASE: DENTAL CARIES.
CC -!- MISCELLANEOUS: GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA
CC 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES
CC WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH
CC FORMS OF GLUCANS.
CC -!- SIMILARITY: TO OTHER GLYCOSYLTRANSFERASES AND SOME TO A GLUCAN-
CC BINDING PROTEIN FROM S.MUTANS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90213; BAA14241.1; -.
DR PIR; A38175; A38175.
DR InterPro; IPR002479; CW_Binding.
DR InterPro; IPR003318; GH_70.
DR Pfam; PF01473; CW_binding_1; 16.
DR Pfam; PF02324; Glyco_hydro_70; 1.
KW Transferrase; Glycosyltransferase; Signal; Repeat; Dental caries.
FT SIGNAL 1 38 POTENTIAL.
FT CHAIN 39 1592 GLYCOSYLTRANSFERASE-I.
FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).
FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).
FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.
FT REPEAT 1093 1142 1.
FT REPEAT 1158 1207 2.


```
DR EMBL; M29154; AAA29646.1; -.
DR EMBL; X56851; CAA40180.1; -.
DR EMBL; S53996; AAD13870.1; -.
DR PIR; A32547; DVZQF.
DR PIR; S18204; S18204.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR003439; ABC_transportr.
DR InterPro; IPR001140; ABCtransprtm.
DR Pfam; PF00005; ABC_tran; 2.
DR Pfam; PF00664; ABC_membrane; 2.
DR ProDom; PD000006; ABC_transportr; 2.
DR SMART; SM00382; AAA; 2.
DR PROSITE; PS00211; ABC_TRANSPORTER; 2.
KW ATP-binding; Glycoprotein; Transmembrane; Transport; Repeat.
FT DOMAIN 1 55 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 56 82 POTENTIAL.
FT TRANSMEM 91 116 POTENTIAL.
FT TRANSMEM 160 188 POTENTIAL.
FT TRANSMEM 194 212 POTENTIAL.
FT TRANSMEM 279 298 POTENTIAL.
FT TRANSMEM 314 338 POTENTIAL.
FT DOMAIN 339 788 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 789 807 POTENTIAL.
FT TRANSMEM 825 846 POTENTIAL.
FT TRANSMEM 908 928 POTENTIAL.
FT TRANSMEM 1028 1048 POTENTIAL.
FT TRANSMEM 1063 1083 POTENTIAL.
FT DOMAIN 1084 1419 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 643 661 POLY-ASN.
FT NP_BIND 413 420 ATP (POTENTIAL).
FT NP_BIND 1161 1168 ATP (POTENTIAL).
FT REPEAT 1 721
FT REPEAT 722 1419
FT CARBOHYD 228 228 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 964 964 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1419 AA; 162251 MW; 0F96C7C1850B33D0 CRC64;

Query Match
Best Local Similarity 21.3%; Pred. No. 89;
Matches 47; Conservative 31; Mismatches 100; Indels 43; Gaps 8;

Qy 27 KOIYYN-----SKAITSSEKSADOFLTNTLLFKGFF---TGHWPYNDLLVLDLST 74
Db 1011 KAIDYKNGQKRRIIVNAALMGFSQSLFINSFAYWFGSLIKRGTTILVDDFMKSLFTF 1070
Qy 75 AATSEYEGSSVDLYG-----AAYGYQCA-----GGTPNKTACMYGGVTLHD 115
Db 1071 IFTGSYAGKLSLKGDSNAKLSFEKYYPLMIRKSNIDVRDDGGIRINKNLKGVKVDIKD 1130
Qy 116 -NNRLTEEEKVPI----NLWIDGKOTTPIDKVKTSKK---EVTVOELDQARHYL---- 163
Db 1131 VNFRIISRPNPVIYKNLSFTCDSKTTAIVGETSGKSTFWNLLRFDYDLKNDHIILND 1190
Qy 164 HGKFGLYNSDSFGGKVGRLIVFHSSEGSTVSVDLFDAGQ 204
Db 1191 MTNFQDYQNNNNNSLVKKNVEFSNQSGSAEDYTVFNNNGE 1231
```

Search completed: January 8, 2003, 11:57:17
Job time : 11.6343 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 8, 2003, 11:54:52 ; Search time 15.4475 Seconds
(without alignments)
3107.877 Million cell updates/sec

Title: US-09-900-766-2
Perfect score: 1218
Sequence: 1 SEKSEINEKDLRKSELQ.....RDNTTISSTLSISLYLVT 233

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mmc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	948	77.8	260	16 Q931M4	Q931M4 staphylococ
2	918	75.4	260	16 Q99SU3	Q99SU3 staphylococ
3	754	61.9	268	2 O85217	O85217 staphylococ
4	420.5	34.5	258	16 Q9EZM4	Q9EZM4 staphylococ
5	372.5	30.6	260	16 Q99T46	Q99T46 staphylococ
6	372.5	30.6	261	2 Q9EZM8	Q9EZM8 staphylococ
7	355	29.1	179	2 Q52075	Q52075 staphylococ
8	311	25.5	240	16 Q9F0L7	Q9F0L7 staphylococ
9	308	25.3	217	2 Q8RR76	Q8RR76 staphylococ
10	308	25.3	241	2 Q53585	Q53585 staphylococ
11	296.5	24.3	242	16 Q85383	Q85383 staphylococ
12	294.5	24.2	225	2 Q8VWV1	Q8VWV1 streptococ
13	293.5	24.1	225	2 Q91921	Q91921 streptococ
14	288.5	23.7	218	2 Q8RR75	Q8RR75 staphylococ
15	288.5	23.7	239	2 O05157	O05157 staphylococ
16	287.5	23.6	225	16 Q99Z21	Q99Z21 streptococ

17	287.5	23.6	239	2 Q06535	Q06535 staphylococ
18	284.5	23.4	239	2 Q06531	Q06531 staphylococ
19	284	23.3	239	2 Q9EZM7	Q9EZM7 staphylococ
20	282	23.2	239	16 Q99T47	Q99T47 staphylococ
21	278.5	22.9	256	2 O8VLM7	O8VLM7 staphylococ
22	277.5	22.8	239	2 Q06532	Q06532 staphylococ
23	277.5	22.8	239	2 Q06533	Q06533 staphylococ
24	276.5	22.7	234	2 Q9R5X4	Q9R5X4 staphylococ
25	276.5	22.7	242	2 Q93CC6	Q93CC6 staphylococ
26	265.5	21.8	271	2 Q9F0L6	Q9F0L6 staphylococ
27	264.5	21.7	239	2 Q53678	Q53678 staphylococ
28	263	21.6	242	2 O54476	O54476 staphylococ
29	263	21.6	242	2 Q93CC5	Q93CC5 staphylococ
30	258.5	21.2	239	2 Q06534	Q06534 staphylococ
31	254.5	20.9	236	2 P97163	P97163 streptococ
32	253.5	20.8	236	2 Q54779	Q54779 streptococ
33	252.5	20.7	222	2 Q9R931	Q9R931 streptococ
34	251.5	20.6	236	2 Q54696	Q54696 streptococ
35	251	20.6	260	2 Q54739	Q54739 streptococ
36	251	20.6	260	2 Q54971	Q54971 streptococ
37	251	20.6	260	2 Q54738	Q54738 streptococ
38	250	20.5	222	2 Q95524	Q95524 streptococ
39	250	20.5	222	2 Q938P4	Q938P4 streptococ
40	249	20.4	258	2 Q9ZNF2	Q9ZNF2 staphylococ
41	248.5	20.4	236	2 Q57453	Q57453 streptococ
42	241	19.8	233	2 Q8RR77	Q8RR77 staphylococ
43	241	19.8	258	2 Q9EZM3	Q9EZM3 staphylococ
44	241	19.8	258	16 Q85382	Q85382 staphylococ
45	238.5	19.6	259	2 Q936G4	Q936G4 staphylococ

ALIGNMENTS

RESULT 1
Q931M4
ID Q931M4 PRELIMINARY; PRT; 260 AA.
AC Q931M4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Enterotoxin P.
GN SEP OR SAV1948.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003364; BAB8110.1;
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin; 1.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 260 AA; 30016 MW; 15CD236270FA8241 CRC64;

Query Match 77.8%; Score 948; DB 16; Length 260;
Best Local Similarity 76.4%; Pred. No. 4.3e-71;
Matches 178; Conservative 21; Mismatches 34; Indels 0; Gaps 0;

```
QY 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNSKAITSEKSADQFLNTLLFKGFFTG 60
Db 28 SEKSEINEKDLRKSELQGTALGNLKOIYYNSKAITSEKSADQFLNTLLFKGFFTD 87
QY 61 HPWYNDLLVDLGSSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLT 120
Db 88 HSWYNDLLVDFDSKDTVDYKYGKKVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLT 147
QY 121 EEKKVPINLWIDGKQTVPIDVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 180
Db 148 EEKKVPINLWIDGKQTVPIETVKTKNKTVOELDLQARHYLHGKFGLYNSDSFGKQV 207
QY 181 RGLIVFHSSEGSTSVSYDLFDAQGOQPDTLRIYRDNTTISSTLSLSLYTT 233
Db 208 RGLIVFHTSTEPSVNYDLFGAOGQSYNTLLRIYRDNTTISENMHIDIYLYTS 260

RESULT 2
Q99SU3
ID Q99SU3 PRELIMINARY; PRT; 260 AA.
AC Q99SU3;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Enterotoxin P.
OS Enterotoxin P.
GN SEP OR SA1761.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,
RA Matsumaru H., Maruyama A., Murakami H., Hoshino A., Mizutani-Uji Y.,
RA Takahashi N.K., Sawano T., Inoue R.I., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF003135; BAB43036.1; -.
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Staph_strep_toxin.
DR Pfam; PF01123; Staph_strep_toxin; 1.
DR Pfam; PF02876; Staph_strep_toxin_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Complete proteome.
SQ SEQUENCE 260 AA; 29708 MW; 087C584EC028CFDB CRC64;

Query Match 75.4%; Score 918; DB 16; Length 260;
Best Local Similarity 73.0%; Pred. No. 1.4e-68;
Matches 170; Conservative 21; Mismatches 42; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELQGTALGNLKOIYYNSKAITSEKSADQFLNTLLFKGFFTG 60
Db 28 SEKSEINEKDLRKSELQGTALGNLKOIYYNSKAITSEKSADQFLNTLLFKGFFTD 87
QY 61 HPWYNDLLVDLGSSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLT 120
Db 88 HSWYNDLLVDFDSKDTVDYKYGKKVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLT 147
QY 121 EEKKVPINLWIDGKQTVPIDVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 180
Db 148 EEKKVPINLWIDGKQTVPLGTVKTKNKTVOELDLQARHYLHGKFGLYNSDSFGKQV 207
QY 181 RGLIVFHSSEGSTSVSYDLFDAQGOQPDTLRIYRDNTTISSTLSLSLYTT 233
Db 208 RGLIVFHTSTEPSVNYDLFGAOGQSYNTLLRIYRDNTTISENMHIDIYLYTT 260
```

```
RESULT 3
O85217
ID O85217 PRELIMINARY; PRT; 268 AA.
AC O85217;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Enterotoxin J.
GN SEJ.
OS Staphylococcus aureus.
OG Plasmid pTB485.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KSI1410;
RX MEDLINE=99052098; PubMed=9835033;
RA Zhang S., Iandolo J.J., Stewart G.C.;
RT "The enterotoxin D plasmid of Staphylococcus aureus encodes a second
RT enterotoxin determinant (sej).";
RL FEMS Microbiol. Lett. 168:227-233(1998).
DR EMBL; AF053140; AAC78590.1; -.
DR HSSP; P13163; 1ESF.
DR InterPro; IPR001961; Staph_strep_toxin.
DR Pfam; PF01123; Staph_strep_toxin; 1.
DR Pfam; PF02876; Staph_strep_toxin_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Plasmid.
SQ SEQUENCE 268 AA; 31230 MW; AC5F3546060ACE22 CRC64;

Query Match 61.9%; Score 754; DB 2; Length 268;
Best Local Similarity 59.3%; Pred. No. 6.3e-55;
Matches 137; Conservative 38; Mismatches 56; Indels 0; Gaps 0;

QY 3 KSEETNEKDLRKSELQGTALGNLKOIYYNSKAITSEKSADQFLNTLLFKGFFTG 62
Db 27 KNETIKENLHKKSELSSITLNNLRHIVFFNEKISEKIMTEDQFLDYLTLFKSFFISHS 86
QY 63 WYNDLLVDLGSSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLT 122
Db 87 QYNDLLVDFDSKDTVDYKYGKKVDLYGAYGYQCAGGTPNKTCMYGGVTLHNNRLT 146
QY 123 KVPINLWIDGKQTVPIDVKTSKKEVTVOELDLQARHYLHGKFGLYNSDSFGKQV 182
Db 147 KVPINLWIDGKQTVPIETVKTKNKTVOELDLQARHYLHGKFGLYNSDSFGKQV 206
QY 183 LIVFHSSEGSTSVSYDLFDAQGOQPDTLRIYRDNTTISSTLSLSLYTT 233
Db 207 LIVFHTSKPELVSYDLFNVIGQYDPKLLKIYQDNKIIESENHIDIYLYTS 257

RESULT 4
Q9EZM4
ID Q9EZM4 PRELIMINARY; PRT; 258 AA.
AC Q9EZM4;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sek (Enterotoxin Sen).
GN SEK OR SEN OR SAV1825 OR SA1643.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699).
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
```

[illegible]

RA	Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA	Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA	Kanamaru M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA	Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA	Sekinizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA	Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA	Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.,
RT	"Whole genome sequencing of methicillin-resistant <i>Staphylococcus aureus</i> ."
RL	Lancet 357:1225-1240(2001).
DR	ENBL; AP003363; BAB57992.1; -;
DR	ENBL; AP003135; BAB42916.1; -;
DR	HSSP; P13163; 1SXT.
DR	InterPro; IPR001961; Strep/Strep_toxin.
DR	Pfam; PF011123; Strep_Strp_toxin; 1.
DR	Pfam; PF02876; Strep_strp_tox_C; 1.
DR	PRINTS; PR00279; BACTRLTOXIN.
DR	PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW	Complete proteome.
SEQ	SEQUENCE 260 AA; 29836 MW; AC927DDCC3E3B8EF CRC64;
Query Match 30.6%; Score 372.5; DB 16; Length 260;	
Best Local Similarity 38.8%; Pred. No. 3.5e-23;	
Matches 93; Conservative 32; Mismatches 92; Indels 23; Gaps	
QY	8 NEKD-----LRKSELOGTALGNLKOITYVNS-KAITSSEKSADQFLTNTLLFGFFETGH 61 : : : : : : : : : : : : : : : :
DB	29 NEEDPKIESLCKSSVDPIAHNINDYINNRTTVKSVSTEKFDELDFLEKSI---- 84 : : : : : : : : : : : : : : : :
QY	62 PW-----YNDLVLDLGSTAATSEYEGSDVLYGAYGYQCAGGTTPNKACMGVGLH 114 : : : : : : : : : : : : : : : :
DB	85 NWLDGISAEFKDLKVFEFSSAISKEFLGKTVDIVGYVKACHGEHQVDFACTYGGVTPH 144 : : : : : : : : : : : : : : : :
QY	115 DNNRLTEKKVPNLWDGQQTPTIDKYKTSKEYTVVOELDQARHYLHGKFGLYNSDS 174 : :
DB	145 ENNKLSPKNIGAVYKDNVNVAFTI--VTTDKKKVTAQELDIKVRTKLNNAYKL--DR 200 :
QY	175 FGGKVQRGVLVFHSSEGSTVS---YDLFDAQGQYPDTLLRIYRDNTTISSTLSLSILYLT 232 :
DB	201 MTSDVQKGVIKFIHSHSEHKESFYDFLYIKGNLPDQYLQIYNDKNTIDSSDYHDIVLET 260 :
RESULT 6	
Q9EZM8	PRELIMINARY; PRT; 261 AA.
AC	Q9EZM8; Q9EZM8;
DT	01-MAR-2001 (TrEMBLrel. 16, Created)
DT	01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	SEO.
GN	SEO.
OS	<i>Staphylococcus aureus</i> .
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC	<i>Staphylococcus</i> .
OX	NCBI_TaxId=1280;
RN	[1]
RC	SEQUENCE FROM N.A.
RP	STRAIN=A900322;
RX	MEDLINE=20571956; PubMed=11123352;
RX	Jarrad S., Peyrat M.A., Lim A., Tristan A., Bes M., Mougel C.,
RT	Etienne J., Vandenesch F., Bonneville M., Lina G.;
RT	"egc, A highly prevalent operon of enterotoxin gene, forms a putative
RT	nursery of superantigens in <i>Staphylococcus aureus</i> ."
RL	J. Immunol. 166:669-677(2001).
DR	ENBL; AF285760; AAC36951.1; -;
DR	HSSP; P13163; 1SXT.
DR	InterPro; IPR001961; Strep/Strep_toxin.
DR	Pfam; PF011123; Strep_Strp_toxin; 1.
DR	Pfam; PF02876; Strep_strp_tox_C; 1.
DR	PRINTS; PR00279; BACTRLTOXIN.
DR	PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.

```

SQ SEQUENCE 261 AA: 29949 MW: 24DA4DDD766288EC CRC64;
Query Match 30.6%; Score 372.5; DB 2; Length 261;
Best Local Similarity 38.8%; Pred. No. 3.5e-23;
Matches 93; Conservative 32; Mismatches 92; Indels 23; Gaps 7;

QY 8 NEKD-----LRKKSELQGTALGNLKIYYNS-KAITSEKSDAQLNTLLFKGFFFTGH 61
   ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 30 NEEDPKIESLCKSSVDPIALHINDDYINNRFVTVKSVISTEXFLDFDLFKSI---- 85

QY 62 PW-----YNDLLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCACMYGGVTLH 114
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 86 NWLDGISAEFKDLKVEFSSAISKEFLGTVDIYGVYKKAHGHEQVDTACTYGGVTPH 145

QY 115 DNNRLTEKKVPINLWIDGKQTTPIDKVKTSKKEVTQVELDQARHYLHGKFGLYNDS 174
   ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 146 ENNKUSEPNKGVAVIKDWNVNTFI--VTDKKKVTQAQELDIKVRTKLNNAYKLY--DR 201

QY 175 FGGKVQRGVLIVFHSSEGSTVS--YDLFDAQGYQPDTLRIYRDNTTISTSLISLYLYT 232
   ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 202 MTSVDQVGKVKFHSSEHKESFYDLFYIKGNLPDQYLOIYNDKNTIDSSDYHIDVYLET 261

RESULT 7
Q52075 PRELIMINARY; PRT: 179 AA.
AC Q52075;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Enterotoxin D.
GE VIRC2.
OS Staphylococcus aureus.
OG Plasmid pIB485.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales.
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89359112; PubMed=2549000;
RA Bayles K.W., Iandolo J.J.;
RT "Genetic and molecular analyses of the gene encoding staphylococcal enterotoxin D.";
RL J. Bacteriol. 171:4799-4806(1989).
DR EMBL; M94872; AAA98133.1; -.
DR HSSP; P13163; 1SXT.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_Strp_toxin; 1.
DR Pfam; PF02876; Strep_Strp_tox_C; 1.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
KW Plasmid.
SQ SEQUENCE 179 AA: 20563 MW: 579FFB811BC08747 CRC64;

Query Match 29.1%; Score 355; DB 2; Length 179;
Best Local Similarity 45.3%; Pred. No. 6.1e-22;
Matches 68; Conservative 26; Mismatches 56; Indels 0; Gaps 0;

QY 1 SEKSEINEKDLRKSELOGTALGNLKIYYNSKAITSEKSDAQLNTLLFKGFFTG 60
   ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 26 NENIDSVREKELHKSELSSTALNNKHYSADKNPTIGENKSTGDOFENTLLYKKFFTD 85

QY 61 HPWYNDDLVDLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTCACMYGGVTLHDNNRLT 120
   ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 86 LINFEDLLINFSNKEMAHFKSKNVDVPIRYSINCYGGEIDRTACTYGGVTPHEGNKLIK 145

QY 121 EKKVPINLWIDGKQTTPIDKVKTSKKEV 150
   ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 146 ERKKIPINLWINGVQKEVSLDKVQTDKKNL 175

RESULT 8
Q9F0L7 PRELIMINARY; PRT: 240 AA.
ID Q9F0L7
```

```

AC Q9F0L7;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Sel (Extracellular enterotoxin L).
GN SEL OR SAV2008 OR SA1816.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales.
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=S.aureus; PubMed=111414901;
RX MEDLINE=20566668;
RA Fitzgerald J.R., Monday S.R., Foster T.J., Bohach G.A., Hartigan P.J.,
RA Meaney W.J., Smyth C.J.;
RT "Characterization of a putative pathogenicity island from bovine
RT Staphylococcus aureus encoding multiple superantigens.";
RL J. Bacteriol. 183:63-70(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Fuyuki K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet. 357:1225-1240(2001).
DR EMBL; AF217235; AAG29598.1; -.
DR EMBL; AP003364; BAB58170.1; -.
DR EMBL; AP003135; BAB43096.1; -.
DR HSSP; P13163; 1ESF.
DR InterPro; IPR001961; Strep/Strep_toxin.
DR Pfam; PF01123; Strep_Strp_toxin; 1.
DR Pfam; PF02876; Strep_Strp_tox_C; 1.
DR PRINTS; PR00277; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Complete proteome.
SQ SEQUENCE 240 AA: 27496 MW: F40D62DA73197881 CRC64;

Query Match 25.5%; Score 311; DB 16; Length 240;
Best Local Similarity 33.3%; Pred. No. 4.1e-18;
Matches 74; Conservative 36; Mismatches 78; Indels 34; Gaps 6;

QY 24 GNLKQIY----YYNSKAITSEKSDAQLNTLLFKGFFTGHPWYND-LLVDLGSTAATS 78
   ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 31 GNLRFYTKYEVNKLNVKDKNSPESHRL-----YSYKNDTLYAEFDNEYITS 79

QY 79 EYEGSSVDLYGAYGYQCAGGTPNKTCACMYGGVTLHDNNRLTEKKVPINLWIDGKQTV 138
   ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 80 DLKGNVDVFGISYK----GSNRT--YGGVTKAENKNKLDSPRIIPINLNGKHQTV 133

QY 139 PLDKVKTSKKEVTQVELDQARHYLHGKFGLY-----NSDSFGGKVGQGLIVF 186
   ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:| ||:|
Db 134 TTKSVSTDKMVTAGEIDVKLRKYLQDEENIYGHNDTCGKEYGTSKFYSGFDKGSVVF 193

QY 187 HSSEGSTVSYDLFDAQGYQPDTLRIYRDNTTISTSLISL 228
   | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db 194 HMDGSGNSFYDLFTYTCYGLPESFLKIYKDNKTVSDTQPHLDV 235

RESULT 9
Q8RR76 PRELIMINARY; PRT: 217 AA.
ID Q8RR76
AC Q8RR76;
```



```
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Enterotoxin H (Fragment).
GN SEH.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21871379; PubMed=11880405;
RA Omoie K., Ishikawa M., Shimoda Y., Hu D.L., Ueda S., Shinagawa K.;
RT "Detection of seg, seh, and sei genes in Staphylococcus aureus
RT Isolates and Determination of the Enterotoxin Productivities of S.
RT aureus Isolates Harboring seg, seh, or sei Genes.";
RL J. Clin. Microbiol. 40:857-862(2002).
DR EMBL; AB060536; BAB85990.1; -.
FT NON_TER
SQ SEQUENCE 217 AA; 25143 MW; A8A44E23F31DE80D CRC64;

Query Match 25.3%; Score 308; DB 2; Length 217;
Best Local Similarity 33.9%; Pred. No. 6.4e-18;
Matches 76; Conservative 46; Mismatches 90; Indels 12; Gaps 7;

QY 10 KDLRKKSELOGTALGNLKOIYYNSKATTSSEKSADQFLNTLLFKGFTGHPWVNDLLV 69
DB 1 EDLHDKSELDTALAN--AYGOYNHPFKENIKSDEISGEKDLIFRN--QGDSG-NDLRV 55

QY 70 DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINL 129
DB 56 KPATADLAQKFNKNVDIYGASFYKCEKISENISECLYGGTTL-NSEKLAQERVIGANV 114

QY 130 WIDGKQTTVPIDKVKTSKKEVTQVQLDQARHLYHGKFGLYNSDFSGKVORGLIVEHSS 189
DB 115 WVDGLOKQETEL--ITNKNKNTLQELDIKIRKILSDKYKIYKDS--EISKGLIEFDKM 169

QY 190 EGSTVSYDLFDAQGOYPTDLLRIYRDNNTTISLS--ISLYLYT 232
DB 170 TPDYSFDIYDLKGENDYEDNKIYEDNKTLSDDISHIDVNLTY 213

RESULT 10
Q53585
ID Q53585 PRELIMINARY; PRT; 241 AA.
AC Q53585;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enterotoxin H precursor.
GN SEH.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=D4508;
RX MEDLINE=95053699; PubMed=7964453;
RA Ren K., Bannan J.D., Pancholi V., Cheung A.L., Robbins J.C.,
RA Fischetti V.A., Zabriskie J.B.;
RT "Characterization and biological properties of a new staphylococcal
RT exotoxin.";
RL J. Exp. Med. 180:1675-1683(1994).
DR EMBL; U11702; AAA19777.1; -.
DR HSP; P13163; IESF.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_tox_C; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal.
FT SIGNAL
FT CHAIN 1 24 POTENTIAL.
FT CHAIN 25 241 ENTEROTOXIN H.
```

```
SQ SEQUENCE 241 AA; 27858 MW; 70F77985877616CE CRC64;

Query Match 25.3%; Score 308; DB 2; Length 241;
Best Local Similarity 33.9%; Pred. No. 7.3e-18;
Matches 76; Conservative 46; Mismatches 90; Indels 12; Gaps 7;

QY 10 KDLRKKSELOGTALGNLKOIYYNSKATTSSEKSADQFLNTLLFKGFTGHPWVNDLLV 69
DB 25 EDLHDKSELDTALAN--AYGOYNHPFKENIKSDEISGEKDLIFRN--QGDSG-NDLRV 79

QY 70 DLGSTAATSEYEGSSVDLYGAYGYQCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINL 129
DB 80 KPATADLAQKFNKNVDIYGASFYKCEKISENISECLYGGTTL-NSEKLAQERVIGANV 138

QY 130 WIDGKQTTVPIDKVKTSKKEVTQVQLDQARHLYHGKFGLYNSDFSGKVORGLIVEHSS 189
DB 139 WVDGLOKQETEL--ITNKNKNTLQELDIKIRKILSDKYKIYKDS--EISKGLIEFDKM 193

QY 190 EGSTVSYDLFDAQGOYPTDLLRIYRDNNTTISLS--ISLYLYT 232
DB 194 TPDYSFDIYDLKGENDYEDNKIYEDNKTLSDDISHIDVNLTY 237

RESULT 11
O85383
ID O85383 PRELIMINARY; PRT; 242 AA.
AC O85383;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Extracellular enterotoxin type I precursor (SEI).
GN SEI OR SAV1828 OR SA1646.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RX SPECIES=S.aureus; STRAIN=FRI445;
RX MEDLINE=98298056; PubMed=9632603;
RA Munson S.H., Tremaine M.T., Betley M.J., Welch R.A.;
RT "Identification and characterization of staphylococcal enterotoxin
RT types G and I from Staphylococcus aureus.";
RL Infect. Immun. 66:3337-3348(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX SPECIES=S.aureus; STRAIN=A900322;
RX MEDLINE=20571956; PubMed=11123352;
RA Jarraud S., Peyrat M.A., Lim A., Tristan A., Bes M., Mougel C.,
RA Etienne J., Vandenesch F., Bonneville M., Lina G.;
RT "egc, a highly prevalent operon of enterotoxin gene, forms a putative
RT nursery of superantigens in Staphylococcus aureus.";
RL J. Immunol. 166:669-677(2001).
RN [3]
RP SEQUENCE FROM N.A.
RX SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Murakami A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF064774; AAC26661.1; -.
DR EMBL; AF285760; AAG36953.1; -.
DR EMBL; AP003363; BAB57990.1; -.
DR EMBL; AP003135; BAB42914.1; -.
DR
```

```
DR HSP; P13163; LSXT.
DR InterPro; IPR001961; Staph_Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin_1.
DR Pfam; PF02876; Staph_Strep_toxin_C; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal; Complete proteome.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 242 ENTEROTOXIN TYPE I.
SQ SEQUENCE 242 AA; 27863 MW; C5C8B4ACE5414A8 CRC64;

Query Match 24.3%; Score 296.5; DB 16; Length 242;
Best Local Similarity 32.4%; Pred. No. 6.7e-17;
Matches 77; Conservative 39; Mismatches 85; Indels 37; Gaps 10;

Qy 8 NEKDLKKSELOG-TALGNLKOIY-----YNSKAITSEKSDQFLTWTLFKGFTGHP 62
Db 17 NIKDL---TYAGDGTGVGNLRFNYTKHYIDILGVTDKNLP-----IANOLEFS---TG-- 64

Qy 63 WYNDLLVDLGSAAATSEVSGSVLDYGYGYOCAGGTPNKTACMYGGVTLHDNNRLTEE 122
Db 65 -TNDLISESNWDELSKPKGKGLDFIGDYNCP-----KSKYMGGATL-SGOYLSA 116

Qy 123 KKVPTNLMDGQTTPIDKVKTSKEVTVOELDLQARHYLHGKFLYNSDSFG-GK--- 178
Db 117 RKIPINLVNGKHKTISTDKIATNKKLVTAQIDVKLRRLQEEYNIYGHNTGKGEYV 176

Qy 179 -----VQRLIVFHSSEGSTSVSYDLFDAQQGYPDTLRLRIYRDNTTISSTLSISL 228
Db 177 YSKSYGFENNGKVLFLHNNEKSFSDLYFGDGLVPFLKIYEDNKTIIESEKFLDV 234

RESULT 12
Q8VVW1 PRELIMINARY; PRT; 225 AA.
AC Q8VVW1:
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Exotoxin I.
GN SPEI.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M1;
RX MEDLINE=21259899; PubMed=11359827;
RA Proft T., Arcus V.L., Handley V., Baker E.N., Fraser J.D.;
RT "Immunological and biochemical characterization of streptococcal
RT pyrogenic exotoxins I and J (SPE-I and SPE-J) from Streptococcus
RT pyogenes".
RL J. Immunol. 166:6711-6719(2001).
DR EMBL; AF438524; AAL31571.1; -.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; UNKNOWN_1.
SQ SEQUENCE 225 AA; 25654 MW; DB193667890ACDAF CRC64;

Query Match 24.2%; Score 294.5; DB 2; Length 225;
Best Local Similarity 32.4%; Pred. No. 8.9e-17;
Matches 72; Conservative 40; Mismatches 85; Indels 25; Gaps 8;

Qy 25 NLKOIY-YNSKAITSEKSDQFLTNTLLFKGFTGHPWYN---DLLVDLGSAAATSEY 80
Db 8 NLRLSYDFTVEVKNGINEGPPF-SGSLFYKNI-----PYGNSIELKVELNSVEKANFF 62

Qy 81 EGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDG---KQTT 137
Db 63 SGKRVDIPTLEYSPPCNSNIKNS---YGGITLSDGNRI-DKKNIPVNIIDGVQOKYSY 118

Qy 138 VPIDKVKTSKEVTVOELDLQARHYLHGKFLGY-----NSDSFGKVGQGLIVFHS 188
Db 119 TDISTVSTDKKEVTIQELDVKSRYYLQRFNIYGYDVKDFGRSSRFQSGFEEGNIIFHL 178

Qy 189 SEGSTSVSYDLFDAQQGYPDTLRLRIYRDNTTISSTLSISL 230
Db 179 NSGERISYNLFDTHGHDRESMLKYSNKTAYSDOLHIDIYL 220

RESULT 14
Q8RR75 PRELIMINARY; PRT; 218 AA.
AC Q8RR75:
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Enterotoxin I (Fragment).
GN SEI.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
```

```
Qy 138 VPIDKVKTSKEVTVOELDLQARHYLHGKFLGY-----NSDSFGKVGQGLIVFHS 188
Db 119 TDISTVSTDKKEVTIQELDVKSRYYLQRFNIYGYDVKDFGRSSRFQSGFEEGNIIFHL 178

Qy 189 SEGSTSVSYDLFDAQQGYPDTLRLRIYRDNTTISSTLSISL 230
Db 179 NSGERISYNLFDTHGHDRESMLKYSNKTAYSDOLHIDIYL 220

RESULT 13
Q9L921 PRELIMINARY; PRT; 225 AA.
AC Q9L921:
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE 25.7 kDa protein.
OS Streptococcus equi.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CF32;
RA Artushin S.C., Timoney J.F., Sheoran A.S.;
RT "Identification and molecular characterization of mitogens from
RT Streptococcus equi.".
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF186180; AAF72808.1; -.
DR HSP; P13163; LSXT.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_toxin_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
SQ SEQUENCE 225 AA; 25684 MW; EB43D47B26FB9F89 CRC64;

Query Match 24.1%; Score 293.5; DB 2; Length 225;
Best Local Similarity 32.4%; Pred. No. 1.1e-16;
Matches 72; Conservative 39; Mismatches 86; Indels 25; Gaps 8;

Qy 25 NLKOIY-YNSKAITSEKSDQFLTNTLLFKGFTGHPWYN---DLLVDLGSAAATSEY 80
Db 8 NLRLSYDFTVEVKNGINEGPPF-SGSLFYKNI-----PYGNSIELKVELNSVEKANFF 62

Qy 81 EGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDG---KQTT 137
Db 63 SGKRVDIPTLEYSPPCNSNIKNS---YGGITLSDGNRI-DKKNIPVNIIDGVQOKYSY 118

Qy 138 VPIDKVKTSKEVTVOELDLQARHYLHGKFLGY-----NSDSFGKVGQGLIVFHS 188
Db 119 TDISTVSTDKKEVTIQELDVKSRYYLQRFNIYGYDVKDFGRSSRFQSGFEEGNIIFHL 178

Qy 189 SEGSTSVSYDLFDAQQGYPDTLRLRIYRDNTTISSTLSISL 230
Db 179 NSGERISYNLFDTHGHDRESMLKYSNKTAYSDOLHIDIYL 220

RESULT 14
Q8RR75 PRELIMINARY; PRT; 218 AA.
AC Q8RR75:
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Enterotoxin I (Fragment).
GN SEI.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
```

```

RN      SEQUENCE FROM N.A.
RP      MEDLINE=21871379; Pubmed=11880405;
RA      Omoe K., Ishikawa M., Shimoda Y., Hu D.L., Ueda S., Shinagawa K.;
RT      "Detection of seg, seh, and sei genes in Staphylococcus aureus
RT      Isolates and Determination of the Enterotoxin Productivities of S.
RT      aureus Isolates Harboring seg, seh, or sei Genes.";
RL      J. Clin. Microbiol. 40:857-862(2002).
DR      EMBL; AB080537; BAB85991.1; -.
FT      NON_TER
SQ      SEQUENCE 218 AA; 24909 MW; 101728FB71BF662 CRC64;

Query Match      23.7%; Score 288.5; DB 2; Length 218;
Best Local Similarity 31.5%; Pred. No. 2.7e-16;
Matches 70; Conservative 39; Mismatches 80; Indels 33; Gaps 8;

QY 23 LGNLKQIY----YVNSKAITSSSEKSDAQFLTNTLLFKGFFTHGHPWYNDLLVDLGSTAATS 78
DB 6 VGNLRFYKHYIDUKGVTDKNLP-----IANQLEFS---TG---TNDLISESNWNDEIS 55
QY 79 EYEGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEEKVPINLWIDGKQTTV 138
DB 56 KFKGKKLDIFGIDYNGPC-----KSKYFEGGATL-SGOYLSARKIPINLWYNGKHKTI 108
QY 139 PIDKVKTSKEVTVOELDQARHYLHGKGLYNSDSFG-GK-----VORGLIVF 186
DB 109 STDKIATNKKLTAQAEIDVKRLRYLOEYNIYGHNNHTGKGYKSKFYGFNGNKVLF 168
QY 187 HSSEGSTVSVDLFDAGQGYPDTLRLRIYRDNNTTISSTLSISL 228
DB 169 HLNNEKSFYDLYFTGDLGVPSFLKIYEDNKIESEKFKHLDV 210

RESULT 15
005157
ID Q05157 PRELIMINARY; PRT; 239 AA.
AC Q05157;
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Type C enterotoxin (Fragment).
OS Staphylococcus intermedius
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1285;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=95-011195;
RA Edwards V.M., Deringer J.R., Callantine S.D., Deobald C.F.,
RA Berger P.H., Kapur V., Stauffacher C.V., Bohach G.A.;
RT "Characterization of the type C enterotoxin (SEC-canine) produced by
RT Staphylococcus intermedius pyoderma isolates.";
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; U91526; AAB50248.1; -.
DR HSSP; P23313; IJCK.
DR InterPro; IPR001961; Stap/Strep_toxin.
DR Pfam; PF01123; Stap_strp_toxin; 1.
DR Pfam; PF02876; Stap_strp_tox.C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
DR NON_TER
SQ SEQUENCE 239 AA; 27536 MW; D660644660DE4191 CRC64;

Query Match      23.7%; Score 288.5; DB 2; Length 239;
Best Local Similarity 34.5%; Pred. No. 3e-16;
Matches 81; Conservative 37; Mismatches 96; Indels 21; Gaps 9;

QY 11 DLKKKSELQALGNLQKIY--YVNSKAITSSSEKSDAQFLTNTLLFKGFFTHGHPWYNDLL 68
DB 10 DLHKSEFTGT-MGNMKCLYDDYYS--ATKVKSDYKFLAHLDIYNISDKLKNYDKVK 65
QY 69 VDLGSTAATSEYEGSSVDLYGAYGYQC-----AGGTPNKTACMYGGVTLHDNNRLTE 121
```

```

DB 66 TELLNEDLAKYKDEVVDVYGSNYVNCYFSSKDNVGVKVTGGTKMYGKITKHEGNHFN 125
QY 122 E--KKYPINLWIDGKQTTVPIDKVKTSKEVTVOELDQARHYLHGKGLYNSDSFGKV 179
DB 126 GNIQNVLRVY-ENKRTTISFD-VQTDKKSVTAQELDIKARNFLLINKNKLYEFNS--SPY 181
QY 180 QRGLIVFHSSEGSTVSVDLFDAGQGYPD--TLRLRIYRDNNTTISSTLSISLYLT 232
DB 182 ETGYIKFIESNDNTFWYDMWPAFDQSKYLMYSDNKTVDKSKVKIEVHLTT 236

RESULT 16
Q99Z21
ID Q99Z21 PRELIMINARY; PRT; 225 AA.
AC Q99Z21;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Streptococcal exotoxin I.
DE SPEI OR SPY1007.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;
RX MEDLINE=21192684; Pubmed=11296296;
RA Ferretti J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,
RA Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).
DR EMBL; AE006546; AAK33906.1; -.
DR HSSP; P13163; ISXT.
DR InterPro; IPR001961; Stap/Strep_toxin.
DR Pfam; PF01123; Stap_strp_toxin; 1.
DR Pfam; PF02876; Stap_strp_tox.C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
DR Complete proteome.
KW SEQUENCE
SQ SEQUENCE 225 AA; 25611 MW; CE0D6736CC11CG04 CRC64;

Query Match      23.6%; Score 287.5; DB 16; Length 225;
Best Local Similarity 32.0%; Pred. No. 3.4e-16;
Matches 71; Conservative 40; Mismatches 86; Indels 25; Gaps 8;

QY 25 NLKQIY-YVNSKAITSSSEKSDAQFLTNTLLFKGFFTHGHPWY--DLLVDLGSTAATSEY 80
DB 8 NLRLYSTYDPTVEVKGINEGPPF-SGSLFYKNI----PYGNSSIELKVELNSVEKANFF 62
QY 81 EGSSVDLYGAYGYOCAGGTPNKTACMYGGVTLHDNNRLTEKKVPINLWIDG---KQTT 137
DB 63 SGKRVDIFLTLEYSPPCNSNIKKNS---YGGITLSDGNRI-DKKNPINWIFIDGVOQKYSY 118
QY 138 VPIDKVKTSKEVTVOELDQARHYLHGKGLY-----NSDSFGKGVQRLIVFHS 188
DB 119 TDISTGSTDKKEVTIQELDVKSRYLQKHFNIFYGFDVKDFGRSSRFQSGFEENIIFHL 178
QY 189 SEGSTVSVDLFDAGQGYPDTLRLRIYRDNNTTISSTLSISLYL 230
DB 179 NSGERISYNLFDTGCHGDRSMLKKYSNDKNTAYSDQLHDIYL 220

RESULT 17
Q06535
ID Q06535 PRELIMINARY; PRT; 239 AA.
AC Q06535;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
```

```

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Enterotoxin type C-3 (SEC3) (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX Staphylococcus.
RN NCBI_TaxID=1280;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=FRI 909.
RX MEDLINE=94011313; PubMed=8406814;
RA Marr J.C., Lyon J.D., Roberson J.R., Luper M., Bohach G.A.;
RT "Characterization of novel type C staphylococcal enterotoxins:
RT biological and evolutionary implications.";
RL Infect. Immun. 61:4254-4262(1993).
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC -!- STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- SIMILARITY: THE DIFFERENT S. AUREUS ENTEROTOXINS AND S. PYOGENES
CC PYROGENIC EXOTOXINS ARE ALL RELATED.
DR EMBL; L13377; AAA26621.1; -.
DR HSP; P23313; LJCK.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strp_toxin; 1.
DR Pfam; PF02876; Staph_Strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Superantigen.
FT NON_TER 1
FT SEQUENCE 239 AA; 27648 MW; CC3CB3B04E4119E0 CRC64;

Query Match 23.6%; Score 287.5; DB 2; Length 239;
Best Local Similarity 33.3%; Pred. No. 3.7e-16;
Matches 78; Conservative 43; Mismatches 94; Indels 19; Gaps 9;

QY 11 DLRRKSELQCTALGNLKOIYYNNSKAITSE--KSADQFLTNTLLFKGFTTGHWPYNDLLV 69
Db 10 DLHKSSEFTGT-MGNMK--YLYDDHYVSATKVKSVKFLAHLDIYNIHKKLNLYDKVKT 66
QY 70 DLGTAATSEYEGSSVDLYGAYGYQC-----AGTTPNKTACMYGGVTLHDNNRLTEE 122
Db 67 ELLNEDLAKNKYDEVVDVYGSNYYVNCYFSSKDNVCKVTGKTCMYGGITKHGHNFDNG 126
QY 123 --KKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELQARHYLHGKFLYNSDFGKVKQ 180
Db 127 NLQNVLIRVY-ENKNTISFE-VQTDKSKVTAQELDIKARFLINKKNLYEFS--SPYE 182
QY 181 RGLVPHSSEGSTSVSYDLFDAQGOYPD--TLIRYRDNTTISSTLSISLYLT 232
Db 183 TGYIKFIESNGNTFWYDMMPAPGDKFDQSKYLMYKNDKRMVDSKSVKIEVHLTT 236

RESULT 18
Q06531 ID Q06531 PRELIMINARY; PRT; 239 AA.
AC Q06531;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE Enterotoxin type C-4 (SEC4446) (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX Staphylococcus.
RN NCBI_TaxID=1280;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=4446;
RX MEDLINE=94011313; PubMed=8406814;
RA Marr J.C., Lyon J.D., Roberson J.R., Luper M., Bohach G.A.;
RT "Characterization of novel type C staphylococcal enterotoxins:
RT biological and evolutionary implications.";
RL Infect. Immun. 61:4254-4262(1993).
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC -!- STAPHYLOCOCCAL FOOD POISONING SYNDROME.

```

```

CC -!- SIMILARITY: THE DIFFERENT S. AUREUS ENTEROTOXINS AND S. PYOGENES
CC PYROGENIC EXOTOXINS ARE ALL RELATED.
DR EMBL; L13374; AAA26618.1; -.
DR HSP; P34071; ISE2.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strp_toxin; 1.
DR Pfam; PF02876; Staph_Strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Superantigen.
FT NON_TER 1
FT SEQUENCE 239 AA; 27612 MW; BB7BD6204731ED24 CRC64;

Query Match 23.4%; Score 284.5; DB 2; Length 239;
Best Local Similarity 32.5%; Pred. No. 6.5e-16;
Matches 76; Conservative 45; Mismatches 94; Indels 19; Gaps 9;

QY 11 DLRRKSELQCTALGNLKOIYYNNSKAITSE--KSADQFLTNTLLFKGFTTGHWPYNDLLV 69
Db 10 DLHKSSEFTGT-MGNMK--YLYDDHYVSATKVKSVKFLAHLDIYNISDKRLKNYDKVKT 66
QY 70 DLGTAATSEYEGSSVDLYGAYGYQC-----AGTTPNKTACMYGGVTLHDNNRLTEE 122
Db 67 ELLNEDLAKNKYDEVVDVYGSNYYVNCYFSSKDNVCKVTGKTCMYGGITKHGHNFDNG 126
QY 123 --KKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELQARHYLHGKFLYNSDFGKVKQ 180
Db 127 NLQNVLIRVY-ENKNTISFE-VQTDKSKVTAQELDIKARFLINKKNLYEFS--SPYE 182
QY 181 RGLVPHSSEGSTSVSYDLFDAQGOYPD--TLIRYRDNTTISSTLSISLYLT 232
Db 183 TGYIKFIESNGNTFWYDMMPAPGDKFDQSKYLMYKNDKTVDSKSVKIEVHLTT 236

RESULT 19
Q09EZM7 ID Q09EZM7 PRELIMINARY; PRT; 239 AA.
AC Q09EZM7;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE SEM.
GN SEM.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OX Staphylococcus.
RN NCBI_TaxID=1280;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN=A900322;
RX MEDLINE=20571956; PubMed=11123352;
RA Jarraud S., Peyrat M.A., Lim A., Tristan A., Bes M., Mougel C.,
RA Etienne J., Vandenesch F., Bonneville M., Lina G.;
RT "egc, A highly prevalent operon of enterotoxin gene, forms a putative
RT nursery of superantigens in Staphylococcus aureus.";
RL J. Immunol. 166:669-677(2001).
DR EMBL; AF285760; AAG36952.1; -.
DR HSP; P13163; LSXT.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strp_toxin; 1.
DR Pfam; PF02876; Staph_Strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT NON_TER 1
FT SEQUENCE 239 AA; 27371 MW; 69240BE23C4028A CRC64;

Query Match 23.3%; Score 284; DB 2; Length 239;
Best Local Similarity 30.4%; Pred. No. 7.2e-16;
Matches 69; Conservative 38; Mismatches 80; Indels 40; Gaps 6;

QY 23 LGNLKOIYYNNSKAIT--SSEKADQFLTNTLLFKGFTTGHWPYNDLLVLDGSTAA---- 76
Db 24 VGVNLRNYGYSPIEDHQSNPENNHLSQLVFS-----MDNSTVTAEFN 70

```

[illegible]

OX	NCBI_TaxID=1280;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=740N;	
RX	MEDLINE=94011313; PubMed=8406814;	
RA	Marr J.C., Lyon J.D., Roberson J.R., Lupher M., Bohach G.A.;	
RT	"Characterization of novel type C staphylococcal enterotoxins:	
RT	biological and evolutionary implications.";	
RL	Infect. Immun. 61:4254-4262(1993).	
CC	-1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION	
CC	STAPHYLOCOCCAL FOOD POISONING SYNDROME.	
CC	-1- SIMILARITY: THE DIFFERENT S. AUREUS ENTEROTOXINS AND S. PYOGENES	
CC	PYROGENIC EXOTOXINS ARE ALL RELATED.	
CC	EMBL; L13375; AAA26619.1; -	
DR	HSSP; P34071; ISE2.	
DR	InterPro; IPR001961; Staph/Strep_toxin.	
DR	Pfam; PF01123; Staph_strep_toxin; 1.	
DR	Pfam; PF02876; Staph_strep_toxin; 1.	
DR	PRINTS; P00279; BACTRLTOXIN.	
DR	PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.	
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.	
KW	Enterotoxin; Toxin; Superantigen.	
FT	NON_TER 1	
FT	1	
QY	SEQUENCE 239 AA; 27642 MW; C77009F46BC8D645 CRC64;	
	Query Match 22.8%; Score 277.5; DB 2; Length 239;	
	Best Local Similarity 32.1%; Pred. No. 2.5e-15;	
	Matches 75; Conservative 45; Mismatches 95; Indels 19; Gaps	
QY	11 DLRRKSELOGTALGNLKQIYYNSKAITSSE-KSAOFLTNTLLFGFTGHPWYNDLLV 69	
DB	10 DLHKSSEFTGT-WDNMK--YLYDDHYVSATKVKSVDKFLAHDLIYNSDKLKNYDKVKT 66	
QY	70 DLGSTAATSVEGSSVDLYGAYGYOC-----AGTPNKTACMYGVTLDHNNRLTEE 122	
DB	67 ELLNEDLAKYKDEVDVYGSYVNCYFSSKDNVKGVGKGTCTMGITKHBGNHFDNG 126	
QY	123 --RKVPINIDWCKQTVPIDKVKTSKKEYVTQELQARHYLHGKGLYNSDSFGKQV 180	
DB	127 NLQNVLRVY-ENKRNTISFE-VQTDKKSVTQAGELDIKARNELINKKNLYEFS--SPYE 182	
QY	181 RGLIVFHSSSGSVSYDLFDAQGYPD--TLRIYRDNTTISTSLISLYLT 232	
DB	183 TGYIKFIENGNFTWYDMMPAPGDKFDQSKYLMYNDKNKTVDSKSVKIEVHLTT 236	
RESULT-23		
Q06533	PRELIMINARY; PRT; 239 AA.	
ID	Q06533	
AC	Q06533;	
DT	01-NOV-1996 (TrEMBLrel. 01, Created)	
DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)	
DE	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)	
DE	Enterotoxin type C (SECCOPELAND) (Fragment).	
OS	Staphylococcus aureus.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;	
OX	Staphylococcus.	
QX	NCBI_TaxID=1280;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=MNCOPELAND;	
RX	MEDLINE=94011313; PubMed=8406814;	
RA	Marr J.C., Lyon J.D., Roberson J.R., Lupher M., Bohach G.A.;	
RT	"Characterization of novel type C staphylococcal enterotoxins:	
RT	biological and evolutionary implications.";	
RL	Infect. Immun. 61:4254-4262(1993).	
CC	-1- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION	
CC	STAPHYLOCOCCAL FOOD POISONING SYNDROME.	
CC	-1- SIMILARITY: THE DIFFERENT S. AUREUS ENTEROTOXINS AND S. PYOGENES	
CC	PYROGENIC EXOTOXINS ARE ALL RELATED.	
EMBL; L13378; AAA26622.1; -		
HSSP; P34071; ISTE.		
InterPro; IPR001961; Staph/Strep_toxin.		


```
Matches 74; Conservative 48; Mismatches 102; Indels 19; Gaps 9;
QY 2 EKSEINEKDLARKSPLOQTALGNLKOIYYNSKAITSSSE-KSAQOFLNTLLFKGFFTG 60
Db 1 ESQPDTPDELHKASKFTG-LWENMKVL--YDRYVSATKVSVDKFLAHLDIYINISDKK 57
QY 61 HPWYNLDLVLGSAATSEYSSVDLYGAYGYQC-----AGTGNPKTACMYGGVTL 113
Db 58 LKNYDKVKTLLNEDLAKKYKDEVDVGSYVNCYFSSKDNVGVKVTGGKTCMYGGITK 117
QY 114 HNNRLTEK--KVPINLWIDGKQTTPIDKVKTSKEYTVQELDLQARHYLHGFLYN 171
Db 118 HEGNFDNGKLNQVLRVY-ENKRNITSE-VQTDKKSQAQELDIKARNFLINKNLYE 175
QY 172 SDSFGKVGORGLIVFHSSSGSVSDVLDFAQGOYPD--TLLRIYRDNNTTISSTLSISLY 229
Db 176 FNS--SPYGTGYIKFIENNGNTFWDMPPAGDKFDQSKYLMYNDKNTVDSKSKVIEVH 233
QY 230 LYT 232
Db 234 LTT 236

RESULT 28
O54476
ID O54476 PRELIMINARY; PRT; 242 AA.
AC O54476;
DT 01-JUN-1998 (TReMBLrel. 06, Created)
DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT 01-WAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enterotoxin.
GN ENT.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=RN4282; TRANSPOSON=TN557;
RX MEDLINE=98385824; PubMed=9720870;
RA Lindsay J.A., Ruzin A., Ross H.F., Kurepina N., Novick R.P.;
RT The gene for toxic shock toxin is carried by a family of mobile
RT pathogenicity islands in Staphylococcus aureus.;
RL Mol. Microbiol. 29:527-543(1998).
DR EMBL: U93688; AAC28968.1; -.
DR HSP: P13163; 1SXT.
DR InterPro: IPR001961; Staph/Strep_toxin.
DR Pfam: PF01123; Staph_Strp_toxin; 1.
DR Pfam: PF02876; Staph_Strp_toxin; 1.
DR PROSITE: PS00278; STAPH_STREP_TOXIN_2; 1.
SQ SEQUENCE 242 AA: 27720 MW; 28CB4F2EB6B66B9 CRC64;

Query Match 21.6%; Score 263; DB 2; Length 242;
Best Local Similarity 30.1%; Pred. No. 4.le-14;
Matches 66; Conservative 35; Mismatches 92; Indels 26; Gaps 7;
QY 23 LGNLKOIYYNSKAITSSSEKSAQOFLNTLLFKGFFTGHPWYNLDLVLGSAATSEYEG 82
Db 29 IDNLRNFTYTKKDFVDLKDVKDNDTPIANOLQF-----SNESY-DLISESKDFNKSFKG 82
QY 83 SSVDLGYGYQCAGGTPNKTACMYGGVTLHDNRLTEKKVPINLWIDGKQTTPIDK 142
Db 83 KKLDFEGISYNGQC-----NTKYIVGVTA-TNEYLDKSRNIPINIWINGNHKTISTNK 135
QY 143 VYTSKKEVTQELDLQARHYLHGKFGLY-----NSDSFGK-----VORGLIVFHSSE 190
Db 136 VSTNKKLVTAQEIYKRLYQEEYNIYHNGTKKGEYGHKSKFYSGFNIGKVTFLHNN 195
QY 191 GSTVSYDLF-DAQGOVPDPTLLRIYRDNNTTISSTLSISL 228
Db 196 NDTFSYDLFTGDDGLPKSFLKIYEDNKTVESEKPHLDV 234

RESULT 30
Q06534
ID Q06534 PRELIMINARY; PRT; 239 AA.
AC Q06534;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Enterotoxin type C (SECOVINE) (Fragment).
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=94011313; PubMed=8406814;
RA Marr J.C., Lyon J.D., Roberson J.R., Luper M., Bohach G.A.;
RT "Characterization of novel type C staphylococcal enterotoxins:
RT biological and evolutionary implications.";
RL Infect. Immun. 61:4254-4262(1993).
CC -!- DISEASE: STAPHYLOCOCCAL ENTEROTOXINS CAUSE THE INTOXICATION
CC STAPHYLOCOCCAL FOOD POISONING SYNDROME.
CC -!- SIMILARITY: THE DIFFERENT S. AUREUS ENTEROTOXINS AND S. PYOGENES
CC PYROGENIC EXOTOXINS ARE ALL RELATED.
DR EMBL: L13379; AAA26623.1; -.
DR HSP: P34071; 1SE2.
```



```
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_strp_toxin; 1.
DR Pfam; PF02876; Staph_strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Enterotoxin; Toxin; Superantigen.
FT NON_TER 1
SQ SEQUENCE 239 AA; 27517 MW; F354742619C8D196 CRC64;

Query Match 21.2%; Score 258.5; DB 2; Length 239;
Best Local Similarity 30.0%; Pred. No. 9.5e-14;
Matches 73; Conservative 48; Mismatches 103; Indels 19; Gaps 9;

QY 2 EKSEETNEKDLRKKSELOGTALGNLKOIYY-YNSKAIT--SSEKADQFLNTLLFKGFFTG 60
Db 1 ESQPDPTDELHAKSKFTG-LMENKVL--YDRVVSATKVKSVKDFLAHDLIYNISDKK 57

QY 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQC-----AGTTPNKTKACTMYGGVTL 113
Db 58 LKNYDKVKTLLNEDLAKKDEVDVYGSNYVNCFSKDNVGVKVTGGKTCMYGGITK 117

QY 114 HDNNRLTEE--KKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYN 171
Db 118 HEGNHFQNGNLQNLIRVY-ENKRNITISFE-VQTDKKSVTQAELDKARSFLINKNLYE 175

QY 172 SDSFGKGVGRLIVFHSSEGSVSYDLFAOQGYPD--TLLRIYRDNTTISSTLSISLY 229
Db 176 FNS--SPYETGYIKTIENNGNTFWDMPPAGDKFDOSKYLMMYNDKNTVDSKSVKIEVH 233

QY 230 LYT 232
Db 234 LTT 236

RESULT 31
P97163 PRELIMINARY; PRT; 236 AA.
AC P97163;
DT 01-MAY-1997 (TremBLrel. 03, Created)
DT 01-MAY-1997 (TremBLrel. 03, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Type A exotoxin precursor (Fragment).
GN SPEA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS156, AND MGAS500;
RX MEDLINE=92044323; PubMed=1940804;
RA Nelson K., Schlievert P.M., Sclander R.K., Musser J.M.;
RT "Characterization and clonal distribution of four alleles of the speA
RT gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
RT Streptococcus pyogenes."
RL J. Exp. Med. 174:1271-1274(1991).
DR EMBL; X61556; CAA43754.1; -
DR EMBL; X61557; CAA43755.1; -
DR EMBL; X61560; CAA43758.1; -
DR EMBL; X61555; CAA43753.1; -
DR EMBL; X61558; CAA43756.1; -
DR EMBL; X61559; CAA43757.1; -
DR EMBL; X61554; CAA43752.1; -
DR HSP; P08095; IB12.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_strp_toxin; 1.
DR Pfam; PF02876; Staph_strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal.
FT NON_TER 1 1
```

```
FT SIGNAL <1 22 POTENTIAL.
FT CHAIN 23 >236 TYPE A EXOTOXIN.
FT NON_TER 236
SQ SEQUENCE 236 AA; 27454 MW; 3FB3F41ABDC1A84 CRC64;

Query Match 20.9%; Score 254.5; DB 2; Length 236;
Best Local Similarity 32.0%; Pred. No. 2e-13;
Matches 74; Conservative 44; Mismatches 94; Indels 19; Gaps 10;

QY 4 SEINEKDLRKKSELOGTALGNLKOIYY-YNSKAIT--SSEKADQFLNTLLFKGFFTG 60
Db 17 SQEVEFAQQDPDSQLRSLVKNLQNIYFLYEGDPVTHENVKSVDDLSDLIYN---VS 73

QY 61 HPWYNDLLVDLGSTAAATSEYEGSSVDLYGAYGYQC--AGTTPNKTKACTMYGGVTLHDNNRL 119
Db 74 GPNYDKLKTALNQEMATLFDKNDVIYGVYHLCYLCEAERSACIYGGVYTHGEGNHL 133

QY 120 TEKKVPINLWIDGKQTTVPIDKVKTSKKEVTVOELDLQARHYLHGKFLYNSDSFGKV 179
Db 134 EIPKIVKVSIDGIO-SLSFD-IETNKMVTAQELDYKRVKLYTDNKQLYTNGP--SKY 189

QY 180 ORGLIVFHSSEGSVSYDLFD---AOGQYPTDLLRIYRDNTTISSTLSI 226
Db 190 ETGYIKTFPKNKESEWFDFPEPEFTQSKY----LMIYKDNETLDSNTSQI 236

RESULT 32
Q54779 PRELIMINARY; PRT; 236 AA.
AC Q54779; Q54613; Q54736; Q54740; Q54741;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TremBLrel. 20, Last annotation update)
DE Type A exotoxin precursor (Fragment).
GN SPEA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MGAS624 AND MGAS158 AND MGAS485 AND MGAS491, AND MGAS495;
RX MEDLINE=92044323; PubMed=1940804;
RA Nelson K., Schlievert P.M., Sclander R.K., Musser J.M.;
RT "Characterization and clonal distribution of four alleles of the speA
RT gene encoding pyrogenic exotoxin A (scarlet fever toxin) in
RT Streptococcus pyogenes."
RL J. Exp. Med. 174:1271-1274(1991).
DR EMBL; X61569; CAA43767.1; -
DR EMBL; X61572; CAA43770.1; -
DR EMBL; X61568; CAA43766.1; -
DR EMBL; X61570; CAA43768.1; -
DR EMBL; X61571; CAA43769.1; -
DR HSP; P08095; IB12.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_strp_toxin; 1.
DR Pfam; PF02876; Staph_strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal.
FT NON_TER 1 1
FT SIGNAL <1 22 POTENTIAL.
FT CHAIN 23 >236 TYPE A EXOTOXIN.
FT NON_TER 236
SQ SEQUENCE 236 AA; 27468 MW; 29DF2AD575623A84 CRC64;

Query Match 20.8%; Score 253.5; DB 2; Length 236;
Best Local Similarity 31.6%; Pred. No. 2.4e-13;
Matches 73; Conservative 45; Mismatches 94; Indels 19; Gaps 10;

QY 4 SEINEKDLRKKSELOGTALGNLKOIYY-YNSKAIT--SSEKADQFLNTLLFKGFFTG 60
Db 17 SQEVEFAQQDPDSQLRSLVKNLQNIYFLYEGDPVTHENVKSVDDLSDLIYN---VS 73
```

```
Db 17 SOEVFAQQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVQDQLSHDLIYN---VS 73
Qy 61 HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYOC-AGGTPNKTKACMYGGVTLHDNNRL 119
Db 74 GNYDKLKTTELKNOEMATLFFKDKNDIYGVYHYHLYCLENASACIYGGVTHNEGNHL 133
Qy 120 TEEKKVPINLWDGKQTTVPIDKVKTSKEVTQVELDQARHYLHGKFGLYNSDSFGGKV 179
Db 134 EIPKKIVKVSIDGQ-SLSFD-IETNKKMVTQAQLDYKVRKYLTDNKQLYTNGP--SKY 189
Qy 180 QRLIVFHSSEGSTVSVDLFD----AQGYPTLLRIYRDNNTTISSTLSI 226
Db 190 ETGYIKFKPKNKESEWDFPEPTQSKY----LMIYKDNETLDSNTSQI 236

RESULT 33
Qy 09R931 PRELIMINARY; PRT; 222 AA.
AC 09R931;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
GN Exotoxin A (Fragment).
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D709;
RX MEDLINE=99137798; PubMed=9952369;
RA Hollingshead S.K., Beall B.;
RT "Genetic linkage of exotoxin alleles and emm gene markers for tissue tropism in group A streptococci."
RL J. Infect. Dis. 179:627-636(1999).
DR EMBL; AF055698; AAD11624.1; -.
DR HSSP; P08095; 1B12.
DR InterPro; IPR001961; Stap/Strep_toxin.
DR Pfam; PF01123; Stap_Strp_toxin; 1.
DR Pfam; PF02876; Stap_Strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
FT NON_TER 1
FT NON_TER 222
FT SEQUENCE 222 AA; 25759 MW; 48BB7ADDCD91FBA3 CRC64;

Query Match 20.7%; Score 252.5; DB 2; Length 222;
Best Local Similarity 32.0%; Pred. No. 2.7e-13;
Matches 73; Conservative 44; Mismatches 92; Indels 19; Gaps 10;

Qy 4 SEINEKDLRKSELOGTAL-GNLKQIY-YNSKAIT-SSEKSADQFLTNTLLFGFTG 60
Db 5 SOEVFAQQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVQDQLSHDLIYN---VS 61
Qy 61 HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYOC-AGGTPNKTKACMYGGVTLHDNNRL 119
Db 62 GNYDKLKTTELKNOEMATLFFKDKNDIYGVYHYHLYCLENASACIYGGVTHNEGNHL 121
Qy 120 TEEKKVPINLWDGKQTTVPIDKVKTSKEVTQVELDQARHYLHGKFGLYNSDSFGGKV 179
Db 122 EIPKKIVKVSIDGQ-SLSFD-IETNKKMVTQAQLDYKVRKYLTDNKQLYTNGP--SKY 177
Qy 180 QRLIVFHSSEGSTVSVDLFD----AQGYPTLLRIYRDNNTTISSTLSI 223
Db 178 ETGYIKFKPKNKESEWDFPEPTQSKY----LMIYKDNETLDSNT 221

RESULT 34
Qy 054696 PRELIMINARY; PRT; 236 AA.
ID 054696
```

```
AC 054696;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Type A exotoxin precursor (Fragment).
GN SPEA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MGAS156;
RX MEDLINE=92044323; PubMed=1940804;
RA Nelson K., Schlievert P.M., Selander R.K., Musser J.M.;
RT "Characterization and clonal distribution of four alleles of the speA gene encoding pyrogenic exotoxin A (scarlet fever toxin) in Streptococcus pyogenes."
RL J. Exp. Med. 174:1271-1274(1991).
DR EMBL; X61573; CAA43771.1; -.
DR HSSP; P08095; 1B12.
DR InterPro; IPR001961; Stap/Strep_toxin.
DR Pfam; PF01123; Stap_Strp_toxin; 1.
DR Pfam; PF02876; Stap_Strp_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00277; STAPH_STREP_TOXIN_1; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal.
FT NON_TER 1
FT SIGNAL <1 22 POTENTIAL.
FT CHAIN 23 >236 TYPE A EXOTOXIN.
FT NON_TER 236 236
FT SEQUENCE 236 AA; 27575 MW; 70F54120E79127DF CRC64;

Query Match 20.6%; Score 251.5; DB 2; Length 236;
Best Local Similarity 31.7%; Pred. No. 3.6e-13;
Matches 72; Conservative 42; Mismatches 102; Indels 11; Gaps 7;

Qy 4 SEINEKDLRKSELOGTAL-GNLKQIY-YNSKAIT-SSEKSADQFLTNTLLFGFTG 60
Db 17 SOEVFAQQDPDPSQLHRSSLVKNLQNIYFLYEGDPVTHENVKSVQDQLSHDLIYN---VS 73
Qy 61 HPWYNDLLVDLGSTAATSEYEGSSVDLYGAYGYOC-AGGTPNKTKACMYGGVTLHDNNRL 119
Db 74 GNYDKLKTTELKNOEMATLFFKDKNDIYGVYHYHLYCLENASACIYGGVTHNEGNHL 133
Qy 120 TEEKKVPINLWDGKQTTVPIDKVKTSKEVTQVELDQARHYLHGKFGLYNSDSFGGKV 179
Db 134 EIPKKIVKVSIDGQ-SLSFD-IETSKMMVTAQLDYKVRKHLTDNKQLYTNGP--SKY 189
Qy 180 QRLIVFHSSEGSTVSVDLFDQAGQYPTLLRIYRDNNTTISSTLSI 226
Db 190 ETGYIKFKISKDETTFWDFPEPFNOVKYLMYKDNETLDSSTSQI 236

RESULT 35
Qy 054739 PRELIMINARY; PRT; 260 AA.
AC 054739; Q54737;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DE 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
GN SUPERANTIGEN SSA.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94222556; PubMed=8168951;
RA Reda K.B., Kapur V., Mollick J.A., Lamphear J.G., Musser J.M., Rich R.R.;
```


D	T	01-MAR-2002 (TrEMBLrel_20, Last annotation update)
D	E	Pyrogenic exotoxin A (Fragment).
G	N	SPEA.
O	S	Streptococcus equisimilis.
O	C	Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
O	X	Streptococcaceae; Streptococcus.
O	C	NCBI_TaxID=119602;
R	N	[1]
R	P	SEQUENCE FROM N.A.
R	C	STRAIN=4951;
A	K	Kalia A., Bessen D.E.;
R	T	"Presence of streptococcal pyrogenic exotoxin A and C genes in human isolates of group G Streptococci";
R	L	Submitted (Aug-2001) to the EMBL/GenBank/DDBJ databases.
D	R	EMBL; AY049745; AAJ06068.1.; -
I	n	InterPro: IPR001961; Strep/Strep toxin.
D	R	pfam; PF01123; Stap_Strp_toxin; 1.
D	F	pam; PF02876; Stap_Strp_tox.C; 1.
D	R	PROSITE; PS00277; STAPH_STREP_TOXIN_1; UNKNOWN_1.
D	R	PROSITE; PS00278; STAPH_STREP_TOXIN_2; UNKNOWN_1.
F	T	NON_TER
F	T	NON_TER
F	T	NON_TER
S	Q	SEQUENCE 222 222 25884 MW; 121f846099z818f8 CRC64;
Query Match 30.5%; Score 250; DB 2; Length 222;		
Best Local Similarity 32.9%; Pred. No. 4.4e-13;		
Matches 74; Conservative 40; Mismatches 99; Indels 12; Gaps		
Qy	4	SEENEKDLKKSLQGTAL-GNLKQIY--YYNSKAITSSESADQLFTWTLFLFKGFFTG 60 : : : : : : : : : :
Db	5	SQEYFAQQDPNPSQLHRSSLVKNLIQTIFYLYEGDVPVHHNVKSYDLLSHDIYN---VS 61 : : : : : : : : :
Qy	61	HPWYNDDLVDLGSAATSEVGSGSDVLGYGVYGQC-AGGTPNKTAGMGCVTLDNNRL 110 : : : : : : : : : : :
Db	62	GLNYDKLKTELKRNEMSTLFKNKVNDIVGEYYYHCYLCKNAKRACIYGGVTNHEGNHL 120 : : : : : : : : : : : : : : : :
Qy	120	TEEKYPINLWDIGKTPTVIDKYTKSKKEYTVQELDQAQRHYLHGKFGLYSNDSFGVKV 170 : : : : : : : : : : : : :
Db	122	EIPKNILVKVISDIGIQ-SLSFD-IETSKMVAQEldYKVRKHITDNQNLTNGP--SKY 171 : : : : : : : : : : : : : : : :
Qy	180	ORGLIVPHSSEGSVSYDLDFDAQOQYDTLLRIYRDNTTI--SSTS 223 : : : : : : : : : : : : : : : :
Db	178	ETGIKIISKDKETFNFDFPFEPFNQVKYLMYIKDNETLDSSTS 222 : : : : : : : : : : : : : : : :
RESULT 40		
Q9ZNF2	ID	Q9ZNF2 PRELIMINARY; PRT; 258 AA.
DC	AC	Q9ZNF2;
DR	DT	01-MAY-1999 (TrEMBLrel_10, Created)
DE	DD	01-MAY-1999 (TrEMBRel_10, Last sequence update)
DN	DT	01-MAR-2002 (TrEMBRel_20, Last annotation update)
GS	GC	Enterotoxin type Gv.
OS	OG	Staphylococcus aureus.
OC	OB	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacilliales;
OX	OC	Staphylococcus.
ON	CB	NCBI_TaxID=1280;
RN	[1]	SEQUENCE FROM N.A.
RP	RM	MEDLINE=N0260630; PubMed=10803494;
RA	AB	Abe J., Ito Y., Onimaru M., Kohsaka T., Takeda T.:
RT	RF	Characterization and distribution of a new enterotoxin-related superantigen produced by Staphylococcus aureus.";
RL	RM	Microbiol. Immunol. 44:79-88(2000).
DR	DM	EMBL; AB016487; BAA36693.1; -
DR	HSB	HSP; P01552; 1SBB.
DR	IPR	InterPro: IPR001961; Stap/Strep toxin.
DR	pam	; pfam; PF01123; Stap_Strp_toxin; 1.
DR	pam	; pfam; PF02876; Stap_Strp_tox.C; 1.
DR	PRINTS	; PRO0279; BACTRUOTOXIN.
DR	PSITE	; PROSITE; PS00278; STAPH-STREP_TOXIN_2; 1.
SQ	SEQUENCE	258 AA; 29839 MW; 1229246D83F4FB77 CRC64;

Qy	4	SEELNEKDLRKRSKELQGTAL-GNLKQIYY-YNSKAIT-SSEKSADQFLTNTLLFKGFFTG	60
Db	17	SOEFAQQDPDPQLHRSSLSVKNLQNIYFLYEGDPVTHENKVSQDQLLSHDIYN--VS	73
Qy	61	HPWYNDLLVDLGGTAATSEYEGSSVDLYGAYGYQC-AGTGNKTKACMYGGVTLHDNNRL	119
Db	74	GPVYDKLKTTELKQEWATLFEKKNVDIYSVEYHLCYLCEAERSACIYGGVTNHEGNIL	133
Qy	120	TEEKVPIINLWIDGKQTPVIDKVKTSKKEVTVQELDLQARHYLHGKFLYNSDSFGVK	179
Db	134	EIPKKIVKVSIDGQ-SLSFD-IETNKKMVTQELDYKRVKYLTONKQLYTNGP--SKY	189
Qy	180	QRGLIVFHSSEGSTVSDLEF---AQGYPTDLLRIYRDNTTISSTLSI	226
Db	190	ETGYIKFIPKNKESFWDPEPEFTQSKY---LMIYKDNETLDSNTSQI	236
RESULT 42			
Q8RR77	ID	Q8RR77 PRELIMINARY; PRT; 233 AA.	
AC	Q8RR77;		
DT	01-JUN-2002 (TReMBLrel. 21, Created)		
DT	01-JUN-2002 (TReMBLrel. 21, Last sequence update)		
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)		
DE	Enterotoxin G (fragment).		
GN	SEG.		
OS	Staphylococcus aureus.		
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;		
OC	Staphylococcus		
OX	NCBI_TaxID=1280;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RA	MEDLINE=21871379; PubMed=11880405;		
RX	Omoe K., Ishikawa M., Shimoda Y., Hu D.L., Ueda S., Shinagawa K.;		
RT	"Detection of seg, seh, and sei genes in Staphylococcus aureus		
RT	isolates and Determination of the Enterotoxin Productivities of S.		
RT	aureus Isolates Harboring seg, seh, or sei Genes.";		
RL	J. Clin. Microbiol. 40:857-862(2002).		
RL	EMBL; AB060535; BAB85989.1; -.		
FT	NON_TER	1	
SQ	SEQUENCE 233 AA; 27040 MW; ECEB5287D63BF60D CRC64;		
Query Match 19.8%; Score 241; DB 2; Length 233;			
Best Local Similarity 28.8%; Pred. No. 2.6e-12;			
Matches 72; Conservative 48; Mismatches 88; Indels 42; Gaps			
Qy	3	KSEINEKDLRKRSKELQGTALGNLQIYY---YNSKAITSSKSAQFLTNTLLFKGFFT	59
Db	5	KLDELNKSVDYKNN--KGT-MGNVMNLYTSPVPEGRGVINSR---QFLSHDLIFP--I	54
Qy	60	GHPWYNDLLVDLGGTAATSEYEGSSVDLYGAYGYQCA-----GGTPNKTACM	107
Db	55	EYKSYNEVKTELENTELANNYKDKVDIFGVPFYFCIIPKSEPDINQFEG-----CCM	109
Qy	108	YGGVTLH--DNNRLTEKKVPINLWIDGKQTPVIDKVKTSKKEVTVQELDLQARHYLHG	165
Db	110	YGGTLFNSSENER---DKLITVQVTTIDNRQSLG-FTTITNKNMVTIQELDYKARHWLTK	164
Qy	166	KFGLYNSDSFGGKQVRGLIVFHSSEGSTVSDLEFDAQGYPTD---LLRIYRDNTTISST	222
Db	165	EKKLYEED--GSAPESGYIKFTEKNNTSFWFDLFPKKELVPFVYKFLIYGDKNVYDSK	222
Qy	223	SLISLYLT 232	
Db	223	SIKVEFLNT 232	
RESULT 43			
Q9EZM3	ID	Q9EZM3 PRELIMINARY; PRT; 258 AA.	
AC	Q9EZM3;		
DT	01-MAR-2001 (TReMBLrel. 16, Created)		
DT	01-MAR-2001 (TReMBLrel. 16, Last sequence update)		

```
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE SEGL29P.
GN Staphylococcus aureus.
OS Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A900322;
RX MEDLINE=20571956; PubMed=11123352;
RA Jarraud S., Peyrat M.A., Lim A., Tristan A., Bes M., Mougel C.,
RA Etienne J., Vandenesch F., Bonneville M., Lina G.;
RT "esc, A highly prevalent operon of enterotoxin gene, forms a putative
RT nursery of superantigens in Staphylococcus aureus.";
RL J. Immunol. 166:569-677(2001).
DR EMBL; AF285760; AAG36957.1; -.
DR HSSP; P01552; 1SBB.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_tox_C; 1.
DR PRINTS; PR00279; BACTRLTOXIN.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
SQ SEQUENCE 258 AA; 29956 MW; A85619E2FE21E3B7 CRC64;

Query Match 19.8%; Score 241; DB 2; Length 258;
Best Local Similarity 28.8%; Pred. No. 3e-12;
Matches 72; Conservative 48; Mismatches 88; Indels 42; Gaps 12;

QY 3 KSEETNEKDLRKSELOCTALGNLKOIYY---YNSKAITSEKSADQFLTNLTFKGFPT 59
Db 30 KLELNKVS DYKNN--KGT-MGNVNNLYTSPVVEGRGVNSR---QFLSHDLIFP---I 79
QY 60 GHPWYNDLLVLGSTAATSEYEGSSVDLYGAYGYQCA-----GGTPNKTACM 107
Db 80 EYKSYNEVKTELENTALENNYKDKKVDIFGVFPYFTCIIPKSEPDINQFEG---CCM 134
QY 108 YGGVTLH--DNNRLTEERKVPINLWDGKQTVPIDKVKTSKEVTVOELDLQARHYLHG 165
Db 135 YGLTFNSENER---DKLITVQVTDNRQSLG--FTITTKNMVYTIQELDYKARHWLTK 189
QY 166 KFGLYNSDFGKVGORGLIVFHSSSEGSVSYVDLFDQAQGYPDT---LLRIYRDNFTTIST 222
Db 190 EKKLYEFD--GSAPESGYIKTEKNNTSFWDLPFKKELVPVPYKFLNIYGDKNKVDK 247
QY 223 SISISLYLYT 232
Db 248 SIKMEVFLNT 257

RESULT 44
ID O85382 PRELIMINARY; PRT; 258 AA.
AC O85382;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Extracellular enterotoxin type G precursor.
GN SEG OR S41824 OR S41842.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158878, 158879, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus; STRAIN=FR1572;
RX MEDLINE=98298056; PubMed=9632603;
RA Munson S.H., Tremaine M.T., Betley M.J., Welch R.A.;
RT "Identification and characterization of staphylococcal enterotoxin
RT types G and I from Staphylococcus aureus.";
RL Infect. Immun. 66:3337-3348(1998).
```

```
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=S.aureus (strain Mu50), and S.aureus (strain N315);
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; AF064773; AAC26660.1; -.
DR EMBL; AP003363; BAB57986.1; -.
DR EMBL; AP003135; BAB42910.1; -.
DR HSSP; P01552; 1SBB.
DR InterPro; IPR001961; Staph/Strep_toxin.
DR Pfam; PF01123; Staph_Strep_toxin; 1.
DR Pfam; PF02876; Staph_Strep_tox_C; 1.
DR PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.
KW Signal; Complete proteome.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 258 ENTEROTOXIN TYPE G.
SQ SEQUENCE 258 AA; 29940 MW; E2982101701D012C CRC64;

Query Match 19.8%; Score 241; DB 16; Length 258;
Best Local Similarity 28.8%; Pred. No. 3e-12;
Matches 72; Conservative 48; Mismatches 88; Indels 42; Gaps 12;

QY 3 KSEETNEKDLRKSELOCTALGNLKOIYY---YNSKAITSEKSADQFLTNLTFKGFPT 59
Db 30 KLELNKVS DYKNN--KGT-MGNVNNLYTSPVVEGRGVNSR---QFLSHDLIFP---I 79
QY 60 GHPWYNDLLVLGSTAATSEYEGSSVDLYGAYGYQCA-----GGTPNKTACM 107
Db 80 EYKSYNEVKTELENTALENNYKDKKVDIFGVFPYFTCIIPKSEPDINQFEG---CCM 134
QY 108 YGGVTLH--DNNRLTEERKVPINLWDGKQTVPIDKVKTSKEVTVOELDLQARHYLHG 165
Db 135 YGLTFNSENER---DKLITVQVTDNRQSLG--FTITTKNMVYTIQELDYKARHWLTK 189
QY 166 KFGLYNSDFGKVGORGLIVFHSSSEGSVSYVDLFDQAQGYPDT---LLRIYRDNFTTIST 222
Db 190 EKKLYEFD--GSAPESGYIKTEKNNTSFWDLPFKKELVPVPYKFLNIYGDKNKVDK 247
QY 223 SISISLYLYT 232
Db 248 SIKMEVFLNT 257

RESULT 45
Q936G4
ID Q936G4 PRELIMINARY; PRT; 259 AA.
AC Q936G4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Enterotoxin-like protein.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M;
RX MEDLINE=95050273; PubMed=7961465;
RA Lin W.S., Cunneen T., Lee C.Y.;
RT "Sequence analysis and molecular characterization of genes required
RT for the biosynthesis of type 1 capsular polysaccharide in
RT Staphylococcus aureus.";
RL J. Bacteriol. 176:7005-7016(1994).
```


RT	"The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation	
RT	Mosaic Structure, and Significant Antigenic Variation.";	
RL	J. Exp. Med. 191:1765-1776(2000).	
DR	EMBL; AF143664; AAF66664.1; -	
DR	HSSP; P13163; 1SXT.	
DR	InterPro; IPR001961; Strep/Strep_toxin.	
DR	Pfam; PF01123; Strep_Strp_toxin; 1.	
DR	Pfam; PF02876; Strep_Strp_tox_C; 1.	
DR	PRINTS; PR00279; BACTRLTOXIN.	
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.	
FT	NON_TER	1
SQ	SEQUENCE	209 AA; 24122 MW; 0CF5D429E1B96EFE CRC64;
Query Match 17.8%; Score 217; DB 2; Length 209;		
Best Local Similarity 29.1%; Pred. No. 2.3e-10;		
Matches 58; Conservative 38; Mismatches 67; Indels 36; Gaps		
QY	50	NTLLFGFFTGHPWYNDLLVD-----LGSTAATSEYEGSSVDLYGAY----- 91
Db	6	NSLLRNIIYSTIVVEYSDTIDFKTSHNLVTKKLDVRDARDFINSEMDEYAANDFKAGDK 65
QY	92	-----GYQCAGGTPNKTCMYGGVTLHDNNRLTEEEKKVPINLWIDGKQTTVPIDK 142
Db	66	IAVFSVPFDWNYLSEG---KVIATYTGGITPYQKTSI--PKNIPVNLWINGKQISVPYNE 120
QY	143	VKTSKKEVTVOELDQARHYLHGKFLGYNDSFGGKVGQRLIVFHSSEGS-TVSYDLFDA 201
Db	121	ISTNKTVTTAQAEIDLKVRKFLIAHQHLYSS--GSSYKSKGLVFHTNDNSDKYSFDFLYT 177
QY	202	QGQYPTDLLRIYRDNNTTIS 220
Db	178	GYRDKESIFKVKYDNKSNF 196
RESULT 50		
ID	Q9LAC6	PRELIMINARY; PRT; 209 AA.
AC	Q9LAC6;	
DT	01-OCT-2000 (TReMBLrel. 15, Created)	
DT	01-OCT-2000 (TReMBLrel. 15, Last sequence update)	
DT	01-JUN-2002 (TReMBLrel. 21, Last annotation update)	
DE	Mitogenic exotoxin Z-20 (Fragment)..	
GN	SMEZ-20.	
OS	Streptococcus pyogenes.	
OC	Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;	
OC	Streptococcaceae; Streptococcus.	
OX	NCBI_TaxID=1314;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	STRAIN=10989;	
RX	MEDLINE=20273982; PubMed=10811869;	
RA	Proft T., Moffatt S.L., Weller K.D., Paterson A., Martin D.,	
RA	Fraser J.D.;	
RT	"The Streptococcal Superantigen SMEZ Exhibits Wide Allelic Variation,	
RT	Mosaic Structure, and Significant Antigenic Variation.";	
RL	J. Exp. Med. 191:1765-1776(2000).	
DR	EMBL; AF143670; AAF66669.1; -	
DR	HSSP; P13163; 1SXT.	
DR	InterPro; IPR001961; Strep/Strep_toxin.	
DR	Pfam; PF01123; Strep_Strp_toxin; 1.	
DR	Pfam; PF02876; Strep_Strp_tox_C; 1.	
DR	PRINTS; PR00279; BACTRLTOXIN.	
DR	PROSITE; PS00278; STAPH_STREP_TOXIN_2; 1.	
FT	NON_TER	1
SQ	SEQUENCE	209 AA; 24100 MW; 53049A11599BEA68 CRC64;
Query Match 17.8%; Score 217; DB 2; Length 209;		
Best Local Similarity 29.1%; Pred. No. 2.3e-10;		
Matches 58; Conservative 38; Mismatches 67; Indels 36; Gaps		
QY	50	NTLLFGFFTGHPWYNDLLVD-----LGSTAATSEYEGSSVDLYGAY----- 91
Db	6	NSLLRNIIYSTIVVEYSDTIDFKTSHNLVTKKLDVRDARDFINSEMDEYAANDFKAGDK 65
RA	Fraser J.D.;	

[illegible]

Search completed: January 8, 2003, 11:58:25
Job time : 19.4475 secs

THIS PAGE BLANK (USP)